

Софийски университет „Св. Климент Охридски“
Философски факултет
Катедра „Библиотекознание, научна информация и културна политика“

Мария Асенова Аврамова

ЛИНГВИСТИЧНА ОСИГУРЕНОСТ НА ЕЛЕКТРОННИТЕ БИБЛИОТЕКИ

ДИСЕРТАЦИЯ

за присъждане на образователна и научна степен „доктор“
по професионално направление 3.5. Обществени комуникации и информационни науки
(Книгознание, библиотекознание, библиография)

Научен ръководител: доц. д-р Милена Миланова

София, 2018

СЪДЪРЖАНИЕ

УВОД	4
Глава 1. ЕЛЕКТРОННИТЕ БИБЛИОТЕКИ В СЪВРЕМЕННОТО ИНФОРМАЦИОННО ПРОСТРАНСТВО	11
1.1 Същност на електронната библиотека – дефиниции и функции.....	13
1.1.1. Електронна библиотека и традиционна библиотека.....	14
1.1.2. Електронна библиотека като информационна система	18
1.2. Типология на електронните библиотеки.....	25
1.3. Концептуални модели на електронна библиотека	28
1.3.1. Референтен модел на електронна библиотека DELOS DLRM.....	29
1.3.2. Информационни модели на електронна библиотека	38
1.3.2.1 CIDOC CRM.....	38
1.3.2.2. FRBR.....	41
1.3.2.3. IFLA LRM.....	44
1.3.2.4. CERIF.....	50
1.4. Оперативна съвместимост на електронните библиотеки	54
1.5. Инструментални средства за създаване на електронни библиотеки.....	56
Глава 2. РАЗВИТИЕ НА ЕЛЕКТРОННАТА БИБЛИОТЕКА	66
2.1. Науки за знанието, повлияли развитието на електронната библиотека	69
2.1.1. Изкуствен интелект	69
2.1.2. Компютърна лингвистика.....	72
2.1.3. Организация на знанието	77
2.2. Съвременни уеб технологии и развитието на електронните библиотеки.....	88
2.2.1. Технологиите на Семантичния уеб.....	91
2.2.2. Онтологиите като средство за представяне на знанията в уеб	95
2.2.2.1. Типове онтологии.....	99
2.2.2.2. Приложения на онтологии в електронна библиотека.....	103
2.2.3. Стандарт SKOS	108
2.3. Създаване на семантични електронни библиотеки.....	115
Глава 3. ЛИНГВИСТИЧНО ОСИГУРЯВАНЕ НА ЕЛЕКТРОННА БИБЛИОТЕКА	122

3.1. Лингвистичното осигуряване като компонент на електронната библиотека...	122
3.2. Информационните езици в състава на лингвистичното осигуряване на електронната библиотека	125
3.2.1. Структура на информационните езици	128
3.2.2. Технология за създаване на лексикално-семантична основа на информационен език	136
3.2.3. Типологизация на информационните езици	142
3.3. Предназначение и специфики на информационните езици в лингвистичното осигуряване на електронната библиотека.....	145
3.3.1. Информационни езици за цялостно представяне на документите	145
3.3.2. Информационни езици за представяне на предмета на документите	151
3.3.3. Информационни езици за многоаспектно представяне на съдържанието на документите	161
3.4. Анализ на резултатите от изследване на използвани лингвистични средства в българските електронни библиотеки.	163
3.5. Модел на лингвистично осигуряване на българска електронна библиотека с универсална тематика	181
ЗАКЛЮЧЕНИЕ	211
СПИСЪК НА ИЗПОЛЗВАНАТА ЛИТЕРАТУРА.....	214
СПИСЪК НА ТАБЛИЦИ, ФИГУРИ И ИЛЮСТРАЦИИ.....	238
ПРИЛОЖЕНИЯ	240
<i>Приложение 1. Йерархия на обектите в IFLA LRM: определения и примери</i>	<i>241</i>
<i>Приложение 2. Добри практики за реализиране на електронни библиотеки</i>	<i>247</i>
<i>Приложение 3. Предимства и ограничения на информационни езици, отразяващи съдържанието на документи.....</i>	<i>269</i>
<i>Приложение 4. Съпоставка на полетата на MARC 21 и елементите на Dublin Core с определители.</i>	<i>271</i>

УВОД

Проблемът за търсенето на информация е един от вечните проблеми на човешкото общество¹. Решението му човечеството е намерило в създаването на библиотеката като универсална система за натрупване, съхраняване, организиране, предоставяне на каталогизирани и систематизирани документи, информация и знание. Днес значителна част от документите и информацията се създават и/или съществуват в дигитална форма. Налага се необходимостта от създаването на специфични информационни системи, предназначени за съхраняването, използването и управлението на дигиталните информационни обекти - електронната библиотека.

Актуалност

Появата и широкото разпространение на електронните библиотеки се свързва с решаване на редица обществени потребности: управление на информационните обекти; осигуряване и управление на достъпа до документи и информация; дълговременното им съхраняване; представяне и запазване на научното и културното наследство; подпомагане на аналитичната работа с документите и информацията; повишаване на ефективността на научната и изследователската дейност, и обучението.

Електронната библиотека представлява специфична информационна система, предназначена за съхраняване и ефективно използване на разнообразни колекции от документи (текст, изображение, аудио, видео и др.), обединени в единна организационна структура и достъп. Основните задачи на електронната библиотека са интеграцията на информационни обекти, предоставянето на достъп, търсене и навигиране. Под „интеграция” се разбира обединяване на разнообразна по вид информация от различни източници (документи и метаданни за тях). Електронната библиотека осигурява търсене основано на комплекс от информационни езици. Навигацията в електронната библиотека означава възможност потребителят да търси и открива необходимата му информация с най-висока степен на пълнота и точност в цялото достъпно информационно съдържание.

Реализирането на задачите на електронната библиотека пряко зависи от комплекса използвани информационни езици – лингвистичното осигуряване на електронната библиотека. Те служат за търсене, преобразуване, идентифициране, структуриране и систематизиране на информационни обекти. Следователно качествата на лингвистичното

¹Yiyu Yao, Yi Zeng, Ning Zhong, Xiangji Huang. Knowledge Retrieval (KR). In: *Proceedings of the 2007 IEEE/WIC/ACM International Conference on Web Intelligence*[online].IEEE Computer Society, Silicon Valley, USA, Nov. 2–5, 2007, p. 729–735. [Viewed 15.04.2018]. Available from: https://www.researchgate.net/publication/4309759_Knowledge_Retrieval_KR

осигуряване определя качеството на цялата електронна библиотека и е важен аспект на изследване на електронната библиотека. Обобщеният опит и теория по създаването и използването на информационни езици би бил полезен за създателите на електронни библиотеки, както и в процеса на обучение на студентите по специалностите Библиотечно-информационни науки и Публични информационни системи. В университетска среда голяма част от текста може да служи за учебен курс по дисциплините „Индексиране на информационни ресурси“, „Информационни системи“, „Информационни процеси“, „Информационна архитектура на електронни библиотеки“.

Електронните библиотеки са нова област с интензивно развитие и с установяващи се традиции в образователен аспект. През последните години в Европа се реализират няколко програми за обучение в областта на електронните библиотеки, например – Университета в Бурос, Швеция предлага в магистърската си програма Библиотечно-информационни науки курс „Електронни библиотеки и информационни услуги“; Университетът в Шефилд, Великобритания обучава студенти в магистърска програма „Управление на електронни библиотеки“; Университетът в Парма, Италия и Университетът в Талин, Естония разработват съвместна магистратура – „Електронни библиотеки“ и др. Основната цел на програмите е да се обучават специалисти по управление на дигитални колекции и услуги.

Цел на изследването е създаване на теоретично обоснован модел на лингвистична осигуреност на електронна библиотека (комплекс от необходими и достатъчни информационни езици за изпълнение на основните задачи и функции на електронните библиотеки).

Обект на изследването се явява електронната библиотека като тип информационна система за съхраняване, обработване, структуриране и систематизиране, разпространяване, достъп и търсене на дигитални документи и информация и предоставяне на услуги.

Предмет на изследването е комплексът от информационни езици за търсене, идентификация и систематизация на информационните обекти, приложими в електронните библиотеки.

Работна хипотеза:

За да изпълнява основните си задачи и функции, електронната библиотека трябва да използва подходящи информационни езици, които да гарантират максимално резултативно търсене, идентификация на информационните обекти и взаимодействие с други електронни библиотеки и автоматизирани системи. Те по отделно изпълняват различни функции: търсене, преобразуване, структуриране, систематизиране, идентифициране. Моделът на

лингвистично осигуряване на електронната библиотека трябва да съдържа необходим и достатъчен комплекс от информационни езици.

Осъществяването на поставената цел се реализира чрез решаването на няколко конкретни задачи:

- 1) Анализ на съвременната информационна среда и дефиниране на мястото, същността и функциите на електронните библиотеки в нея;
- 2) Анализ на съвременните концепции за електронна библиотека и проучване на концептуални модели на електронна библиотеки.
- 3) Проучване на добри практики за реализации на електронните библиотеки.
- 4) Проучване на съвременни технологии за развитието на електронните библиотеки и изграждане на семантични електронни библиотеки.
- 5) Изследване и описание на лингвистичните средства, използвани при създаване на съвременни електронни библиотеки.
- 6) Представяне ролята и предназначението на информационните езици в състава на лингвистичното осигуряване на електронната библиотека.
- 7) Провеждане на емпирично изследване на използваните информационни езици в българските електронни библиотеки.
- 8) Разработване на модел на лингвистично осигуряване на българска електронна библиотека с универсална тематика.

Теории и подходи, на които се опира изследването:

Методологически изследването се основава на **системния подход** при изследване обекта на дисертацията – електронните библиотеки, и при разглеждането на информационните езици по отделно и в комплекс като лингвистично осигуряване на електронна библиотека. Изследването на информационните езици се основава на теория на информационното търсене, чиито терминологичен апарат се използва в изследването.

Изследователските задачи са постигнати с използването на качествени и количествени **методи**, формиращи комплексен методичен подход за решаване на един интердисциплинарен изследователски проблем, какъвто е създаването на лингвистично осигуряване на електронната библиотека. Прилагат се методът на теоретичния анализ и синтез, контент анализ, сравнителен анализ, методи за оценка. За представянето на лингвистичните средства са използвани дескриптивен метод на анализ като са приложени наблюдение, идентификация, диференциация и типологизация и сравнителен анализ. Наред с тях, се използват и математически и статистически методи за анализ на използваните лингвистични средства в българските електронни библиотеки. Информацията, необходима

за постигане на целите и задачите на дисертационния труд, се осигурява от теоретично изследване на публикации на български и чуждестранни автори; първични данни за електронните библиотеки, събрани чрез наблюдение в интернет среда.

Структурата на изследването следва естествената логика на темата.

Глава 1 е посветена на електронната библиотека като част от съвременното информационното пространство, в което най-важните субекти продължават да бъдат библиотеките и др. институции, ангажирани в събирането, съхраняването, организирането и предоставянето на достъп до документи, информация и знания. Същността на понятието „електронна библиотека“ се проследява чрез два подхода. При първия електронната библиотека се свързва с технологията на традиционната библиотека и се определя с нейните базисни характеристики: съхраняване и предоставяне на потребителите на достъп до документи и информация за тях, гарантиране на качество и политика на обслужване. В рамките на този подход електронната библиотека може да се разглежда като еволюционен етап от развитието на библиотеката: библиотека с фишови каталози, библиотека с автоматизиран каталог и електронна библиотека. В съпоставка с традиционната библиотека се извеждат предимствата на електронната библиотека: едновременен достъп на множество потребители до един и същи документ, предоставяне на нови форми за работа с документите, бързо обновяване на информацията, достъп навсякъде и по всяко време, развити средства и методи за търсене. Вторият подход при разглеждане на същността на електронната библиотека акцентира върху нейната реализация като автоматизирана информационна система. Редица изследователи следват този подход и дават свои дефиниции. Въпреки това няма общоприето определение на понятието „електронната библиотека“. За целите на изследването бе изведена собствена дефиниция, която определя електронната библиотека като информационна система за надеждно съхраняване и ефективно използване на разнообразни документи (текст, видео, аудио и др.) с унифициран подход към тяхната организация и достъп.

Определението предполага информационна система с архитектура и програмна среда, организация (или лице), която извършва подбор, съхранение и т.н. и гарантира качество на информацията и услугите, политика за използване на електронната библиотека и потребители. Основните функции и задачи на електронната библиотека са интеграцията на информационни обекти (документи и метаданни за тях), предоставянето на достъп, и осигуряване на ефективни средства и методи за търсене и навигиране.

Днес електронната е популярна информационна система с множество и най-разнообразни реализации, което затруднява нейното изследване. Затова предлагаме

типология на съществуващите електронни библиотеки по различни признаци: начин на създаване, вид на документите, според целевото и предназначение, според тематичния обхват, начин на достъп и собственост/принадлежност.

Изграждането на електронна библиотека на съвременния етап на развитие на информационните и комуникационните технологии предполага възможност за широко използване и за взаимодействие (обмен и интеграция на данни) с други информационни системи (електронни библиотеки). Познаването и съобразяването с Референтен модел на електронна библиотека DELOS DLRM гарантира оперативна съвместимост на електронната библиотека. Електронната библиотека трябва да бъде изградена на подходящ информационен модел. Като такива са разгледани моделите: за международен обмен на информация в областта на културното наследство CIDOC CRM, за функционалност на библиографските данни фамилията модели FRBR и IFLA LRM, и за научноизследователската сфера – CERIF.

Специално внимание се отделя на проблема за осигуряване на оперативна съвместимост между електронните библиотеки, особено в аспекта ѝ семантична оперативна съвместимост, който е свързан с интеграцията на документи и информация и точността на информационното търсене. В света са се наложили различни инструментални средства за създаване на електронната библиотека. На база проучване в Директорията на хранилищата с отворен достъп OpenDOAR предлагаме подходящи програмни продукти за създаване на електронна библиотека с отворен код, които са разпространени в световен мащаб и/или са познати в нашата страна – DSpace, EPrints, Greenstone и Invenio.

В глава 2 се разглежда развитието на електронната библиотека като информационна система, основана на знания. Това означава, че е възможно информационната система да взаимодейства с потребителя на естествен език, да „разбира” и използва знания, да прави логически изводи, да предлага подобро търсене и извличане на знания. Тези усъвършенствани информационни системи се създават на базата на най-новите достижения на науките за знанията – изкуствен интелект, компютърна лингвистика и организация на знанията. Приносът на изкуствения интелект в развитието на информационните системи, вкл. и на електронни библиотеки, касае създаването на модели за формализирано представяне и използване на знания и на методи за тяхната машинна обработка. Компютърната лингвистика подпомага решаването на редица задачи, свързани с автоматичната обработка на естествения език, като изгражда лексикални семантични ресурси. Те могат да бъдат използвани за концептуално индексване и търсене в електронните библиотеки. Организацията на знанието е тясно свързана с библиотечно-

информационните науки, като в тесния си смисъл има най-изявена връзка с класификациите на информационните ресурси, тяхното описание и индексирание. Системите за организация на знания (като напр. речници, нормативни файлове, тезауруси, предметни рубрики, класификации и мн. др.) улесняват управлението и търсенето на информация и знания от хора и машини. Специално е разгледано тяхното разнообразие и предназначение в създадена за целта типология.

Като съществен фактор за развитието на електронната библиотека като система функционираща в метасредата на интернет се разглежда технологичната еволюция на уеб. Съвременният етап на уеб, нар. Семантичен уеб се характеризира със семантични технологии; средства от областта на изкуствения интелект, компютърната лингвистика и организацията на знанията; формати за представяне на данни, които гарантират многократно използваемост за различни приложения; разработване на технологии за комуникация между приложенията в интернет; осигуряване на лесен достъп до информация на основа свързаните данни и др. Представени са най-новите технологии, на които се изгражда концепцията за семантичния уеб.

Разгледана е същността на онтолозиите като базово средство за семантично представяне на знанията в уеб и е представено тяхното типово разнообразие. Специално внимание се отделя на приложенията на онтолозиите в електронна библиотека.

Машинноинтерпретируемото представяне на речниците на системите за организация на знания се основава на специално създадени уеб стандарти. По-подробно се разглежда стандартът SKOS за формализирано представяне на различни системи за организация на знания като тезауруси, предметни рубрики, класификации, таксономии, нормативни речници и др. с цел тяхното използване като свързани данни в уеб.

В резултат на развитието на уеб технологиите, електронните библиотеки се усъвършенстват в т. нар. семантични електронни библиотеки. Семантични електронни библиотеки се изграждат на базата на системите за организация на знанията на инструментите на семантичния уеб за изразяване на семантика и за осигуряване на оперативна съвместимост между информационните системи с цел обмен и многократно използване на данни ина решенията на социалния уеб за взаимодействие между потребителите. Разгледани са няколко проекта за създаване на семантични електронни библиотеки.

Глава 3 се спира върху проблемите на лингвистичното осигуряване на електронна библиотека. Представени са разнообразни подходи и теории при дефинирането на понятието, произтичащи от различни научни дисциплини и теории, които определят и

обосновават различни възгледи и способности за изследване на тематиката и предлагащи различни решения за създаване и оптимизиране на процесите в електронната библиотека. Акцентът е поставен върху „класическият подход“, при който се изследват информационните езици за представяне на съдържанието и характеристиките на информационните обекти в електронните библиотеки и свързаните с тях модели за информационно търсене. Представена е структурата на информационните езици: азбука, лексика, отношения, граматика както и технологията на създаването им. Информационните езици са типологизирани по определени критерии и е предложена класификационна система, разгледани са предназначението и спецификите на различните видове информационни езици: за цялостно представяне на документите; за представяне предмета на документа; за многоаспектно представяне на съдържанието на документа. Втората част на глава 3 представя резултатите от проведено наблюдение върху лингвистичното осигуряване на български електронни библиотеки. За целта и въз основа на приетите в изследването определения и ограничения е разработена методика и са определени критерии за провеждането на емпиричното изследване. Резултатите са обобщени и са използвани при създаването на модел за лингвистично осигуряване на българска електронна библиотека с универсална тематика. Моделът е визуализиран схематично и съдържа необходим и достатъчен комплекс от информационни езици, които да осигурят максимални възможности за търсене а) по всички необходими за търсенето и идентифицирането елементи на библиографското описание и б) по съдържанието на документите. Моделът е конструиран въз основа на теоретичния анализ на информационните езици като специални изкуствени езици, определяне ролята на отделните езици в лингвистичното осигуряване на електронна библиотека, извеждане на предимствата и ограниченията на информационните езици, анализът на резултатите от проведено изследване. При конструирането на модела са следвани основни принципи: международна практика, национална практика, наличие на стандартни информационни езици, максимално осигуряване на разнообразни възможности за търсене. Моделът позволява лингвистичното осигуряване на електронна библиотека въз основа на стандартизирани подходи и инструменти, които водят до планирането, изграждането и реализирането на качествена информационна система, която позволява множество изходи в помощ на потребителя и гарантираща достъп до съдържанието на системата чрез разнообразни нормативно осигурени решения.

Глава 1. ЕЛЕКТРОННИТЕ БИБЛИОТЕКИ В СЪВРЕМЕННОТО ИНФОРМАЦИОННО ПРОСТРАНСТВО

Последните петдесет години се отличават с изключително бързото развитие на информатизацията (компютъризацията)* на обществото. Китайската изследователка Джорджет Уанг описва „информатизацията” като процес на промяна, който включва: а) използването на информацията и информационните технологии до такава степен, че да станат доминиращ ресурс в общественото управление и б) безпрецедентно нарастване на обема и скоростта на произвеждане на информация². Процесът на информатизация е всеобщ и променя напълно облика на съвременната цивилизация – в резултат се формира т.нар. „информационно общество”. Информацията става едновременно преобладаващият и най-значимият ресурс, който поема централната роля на човешкия труд и капитала – основните ресурси през индустриалната епоха.

Напредъкът в информационните и комуникационните технологии въвежда нови „стандарти” в производството, организацията, обработването, разпространението и използването на информация в електронното пространство. Информацията днес се разпространява, преобразува и консумира все повече в електронна среда и чрез най-популярното приложение на интернет – World Wide Web, наричано кратко уеб.

Основната концепция за изграждане на информационното общество – свобода и право на достъп до информация и знания без пространствени, времеви, социални или други ограничения, предполага формиране на единно електронно информационно пространство, чиито градивни елементи са:

- електронни информационни ресурси, съдържащи данни, информация и знания, създадени в резултат на дигитализация или първоначално възникнали в дигитална форма;
- организационни структури, чиито функции са свързани със събиране, обработка, съхраняване, разпространение, търсене и предаване на информация;

* Терминът „информатизация“ е въведен от влиятелните френски общественици Simon Nora и Alain Minc с публикацията им от 1978 г. „L'Informatisation de la Société: Rapport à M. le Président de la République” [Nora, Simon, Minc, Alain. L'informatisation de la société: Rapport à M. le Président de la République. Paris: Seuil, 1978, 162 p. ISBN 2-0-2004974-0], преведена на английски език през 1980 г. като „The computerization of society” [Nora, Simon, Minc, Alain. The computerization of society: A report to the President of France. Cambridge, Mass.: The MIT Press, 1980. p. 208. ISBN 0-26-264020-1], въпреки това авторите и в следващи публикации предпочитат да използват информатизация, а не компютъризация.

² Wang, Georgette. Many Paths to Many Destinations. In: Wang, Georgetteed. *Treading different paths: informatization in Asian nations*. Norwood, N.J.: Ablex Pub. Corp., 1994, p. 247–258. ISBN 1-567500-47-1

– средства за информационно взаимодействие, които осигуряват достъп до информационните ресурси чрез съответстващи информационни и комуникационни технологии.

Най-отличимите участници в електронното информационно пространство са:

– библиотеки и институции за научна и техническа информация, традиционно ангажирани да предоставят качествени информационни ресурси за обслужване на своите читатели/потребители;

– научни и образователни организации - създават и поддържат информационни ресурси (електронни архиви, колекции или образователни портали и ресурси, създавани в рамките на електронното обучение и др. п.);

– обществени организации, които публикуват в интернет – съюзи, общества, асоциации, фондации и други неправителствени организации, заинтересовани в координацията и кооперацията при изпълнение на проекти, свързани с привличане на знания и формирането на електронни пълнотекстови колекции;

– издателства и редакции на периодични издания - създават електронни колекции или архиви, за да предоставят достъп на потребителите до електронни издания при условия, отчитащи интересите на издатели и автори;

– архиви - разширяват условията за достъп до своите специфични фондове (документи на държавните органи и обществени организации, документалното наследство на значими личности и документи за историята на страната), като се съобразят със спецификата на предоставянето на архивни материали на ползвателите;

– музеи – във връзка с традиционна си роля за опазване на културното и историческото наследство разработват културно-просветни и научно-информационни проекти с обща цел – представяне на фондовете си в електронното пространство.

Към представените субекти могат да бъдат добавени и други, но най-важните институционални форми, исторически доказали себе си в традиционното информационно пространство, представляват стабилизиращи фактори и в електронната среда. Според някои автори централна роля в информационното пространство имат библиотеките, не само като пазители на знанията, но и като институция, натрупала опит с тяхното организиране и осигуряване на достъпа до тях. По аналогия типът автоматизирана информационна система за комплектуване, съхраняване, търсене и достъп до дигитални документи е определен като електронна библиотека.

Тук трябва да подчертаем, че се използват различни синоними на понятието „електронна библиотека”: „дигитална библиотека”, „цифрова библиотека”, „виртуална

библиотека”, „онлайн библиотека” и др. В професионалната среда на библиотекарите по-широко разпространеният термин е „дигитална библиотека”, но въпреки това в настоящото изследване предпочитаме „електронна библиотека”. Имаме предвид най-вече характера на предоставяните услуги за потребителите – от разстояние чрез използването на електронни средства. Други наши съображения: а) „електронна библиотека” кореспондира с други библиотечни термини: „електронен каталог”, „електронни издания”, „електронни информационни ресурси”; б) „електронна библиотека” се свързва и с други термини, отразяващи глобални социално-икономически тенденции в развитието на информационното общество: електронното правителство (eGovernment), електронната демокрация (eDemocracy), електронната икономика (e-Economy), електронната наука (eScience) и т.н

1.1 Същност на електронната библиотека – дефиниции и функции

Възникването на електронните библиотеки се свързва със средата на 80-те години на XX век, когато компютърната техника и информационните технологии вече са в състояние да осигурят надеждно съхранение, оперативна обработка и ефективно използване на големи масиви от разнородна информация (текст, аудио, видео, и др.). В края на века се създават сериозни проекти за електронни библиотеки. Въпреки натрупания опит все още съществуват проблеми с еднозначното интерпретиране и дефиниране на понятието „електронна библиотека”³. Отсъствието на общоприето определение се свързва с неясните физически форми и интензивното развитие на този „феномен”⁴. Коментираното понятие, по мнението на специалистите, е „широко” и „разнопланово”⁵. От една страна, то е свързано с технологията на класическата библиотека, а от друга - предполага електронна среда за контролирано съхраняване, обработване и осигуряване на достъп до организирани (структурирани и систематизирани) дигитални документи. Електронната библиотека обаче не може да бъде характеризирана само чрез базисните функции (събиране, организиране, съхраняване и достъп⁶) на библиотеката като институция. Тя има редица свои специфики, които ще бъдат изтъкнати в тази глава.

³ Харизанова, Оля. *Библиотеките и мрежовото общество: Ефекти и трансформации*. София: Унив.изд. „Св. Климент Охридски“, 2010, с. 230.

⁴ Земсков, А. И., Шрайберг, Я. Л. *Электронная информация и электронные ресурсы*. Москва: Фаир, 2007, с. 332.

⁵ Пак там.

⁶ Закон за обществените библиотеки [онлайн]. В: *Lex.bg*. [Прегледан на 12.12.2016]. Достъпно от: <http://www.lex.bg/bg/laws/ldoc/2135636021>

В литературата, посветена на същността на електронната библиотека, се очертават два аспекта на разглеждане, без те да си противоречат и взаимоотричат. Първият се определя като „библиотековедски”, защото свързва феномена „електронна библиотека” с развитието на библиотеката като институция. Вторият разглежда електронната библиотека като тип автоматизирана информационна система, която може да не бъде свързана с традиционната библиотека.

1.1.1. Електронна библиотека и традиционна библиотека

Анализирайки световния опит по създаването и развитието на електронните библиотеки, Е. Горний и К. Вигурский обобщават най-разпространените схващания за формата на съществуване на електронната библиотека в три групи:

1. Бази от данни в интернет.
2. Колекция от информационни ресурси в интернет.
3. Автоматизация на традиционните библиотеки⁷.

Подобно обобщено представяне, естествено, не е прецизно, но подпомага изясняването на същността на електронната библиотека. Разглеждането на електронните библиотеки като бази от данни с достъп чрез интернет съществено стеснява понятието и фактически дава ново название на вече известна същност. Свеждането на електронната библиотека до колекция/и от електронни ресурси лишава понятието от неговия изконен атрибут - системността. Да поясним чрез аналогия – не всяка съвкупност от книги можем да разглеждаме като библиотека. Библиотеката предполага стратегия за формиране и развитие на фонда, каталогизация на информационните обекти и ред други специфики. Несъмнено всяка електронна библиотека е колекция или съвкупност от колекции, но обратното не е вярно – не всяка колекция е библиотека, вкл. и електронна.

Автоматизацията на традиционните библиотеки представлява внедряването на техника, програмни продукти, икономико-математически модели и системи за управление, които частично или напълно освобождават човека от изпълнението на трудоемки рутинни операции и оптимизират процесите по събиране, обработване, съхраняване, преобразуване, предаване и използване на информация.

⁷ Горний, Евгений, Вигурский, Константин. *Развитие электронных библиотек: Мировой и российский опыт, проблемы, перспективы* [онлайн]. PDF, 294.19 KB. [Прегледан 12.12.2016]. Достъпно от: <http://www.ifap.ru/library/book004.pdf>

Електронната библиотека може да бъде разглеждана като етап от еволюционното развитие на класическата. Според израелските библиотековеди Т. Шарон и А. Франк⁸ библиотеката съществува като:

- традиционна (аналогова библиотека, класическа) - с печатен фонд и фишови каталози;
- автоматизирана – традиционна библиотека с електронен каталог;
- електронна – автоматизирана библиотека, в която информацията и документите са в електронен формат⁹.

Определение на понятието „електронна библиотека”, съобразено с „генетичната”, понятийната и функционалната му взаимосвързаност с традиционната библиотека предлага Т. Ф. Берестова:

– Електронната библиотека е фонд от структурирани локални и/или разпределени електронни документи, които чрез информационно-търсещ език или чрез комплекс от езици са свързани в единна идеология на структурализация и достъп, сформирана от библиотечен персонал на основата на материално-техническите възможности и съответстваща на информационните потребности на ползвателите¹⁰.

– Класическата библиотеката, изхождайки от системния подход, се интерпретира като социално-икономическа система, състояща се от подсистемите: библиотечен фонд, читатели, библиотечен персонал и материално-техническа база. Този модел, според неговия създател Ю. Столяров, е валиден и за електронната библиотека. Той я разглежда като „система за достъп”¹¹, предполагаща следните четири компонента: електронни ресурси, организация за индексирание, навигация и търсене, наличие на квалифициран библиотечен персонал, ползватели¹² - тоест всичките четири елемента на класическата библиотека са налице.

⁸ Шарон, Т., Франк, А. Електронные библиотеки в Интернете: Доклад 029-142 на 66-й Генеральной конференции ИФЛА, 13–18 авг. 2000 г., Иерусалим, Израиль. В: *Научные и технические библиотеки*, 2001, № 2, с. 129.

⁹ Шарон, Т., Франк, А. Електронные библиотеки в Интернете: Доклад 029-142 на 66-й Генеральной конференции ИФЛА, 13–18 авг. 2000 г., Иерусалим, Израиль. В: *Научные и технические библиотеки*, 2001, № 2, с. 129.

¹⁰ Берестова, Т. Ф. Электронная библиотека как инструмент создания единого информационного пространства. В: *Вестник Челябинской государственной академии культуры и искусств* [онлайн], 2009, 4, с. 20. [Прегледан 24.05.2018]. Достъпно от: <https://cyberleninka.ru/article/n/elektronnaya-biblioteka-kak-instrument-sozdaniya-edinogo-informatsionnogo-prostranstva> .

¹¹ Земсков, А. И., Шрайберг, Я. Л. *Электронная информация ...*, 2001, с. 7.

¹² Пак там, с. 21.

Библиотечният персонал е съществен елемент на всяка библиотека, в това число и на електронната. Той гарантира качество на информацията, удобни средства за търсене, условия за съхраняване и достъпност. „Библиотекарят“ осигурява подбора на материалите и създава тематичната структурираност и систематизираност на колекциите. Той е посредник и „преводач“ от естествения език на езика за информационно търсене. Без такава „интелектуална обработка“ не може да се говори за електронна библиотека¹³. Ролята на „библиотекаря“ често се пренебрегва от създателите на електронните библиотеки, които най-често имат ИКТ квалификация. За разлика от тях библиотечно-информационните специалисти твърдят, че подборът, организацията и съхранението на информационните ресурси са от основно значение за понятието „библиотека“¹⁴.

Изследователите Т. Шарон и А. Франк представят шест базисни характеристики на всяка библиотека¹⁵:

– *Обекти* - Библиотеката създава съвкупност от „обекти“ (документи), т.е. фонд от книги, периодични издания, електронни и мултимедийни документи и т. н. Обектите трябва да бъдат достъпни непосредствено в библиотеката или чрез мрежа.

– *Метаданни* - Библиотеката обичайно създава метаданни: каталози, ръководства, речници, тезауруси, указатели, реферати, обзори и др.п.

– *Услуги* - Сред услугите се поставят различните средства за достъп (търсене, преглеждане и др.) с оглед категориите читатели и техните интереси, способите за диференцирано разпространение на информация, статистиката и управлението на библиотеката, оценките на ефективността и др.

– *Интереси* - Всяка библиотека има определена област на интереси, в съответствие с които формира своя фонд - напр изкуство, наука или литература. Библиотеката се създава за определена категория потребители: учени, студенти, широк кръг читатели и пр.

– *Контрол на качеството* - Библиотеката предполага наличие на показатели за качество, свързани с подбора на документите в съответствие с профила и предназначението на библиотеката, достъпа до ресурсите чрез създаване на библиографски записи, указатели, реферати и т.н.

¹³ Земсков, А. И., Шрайберг, Я. Л. *Электронная информация...*, с. 346.

¹⁴ Виж в: Witten, I. H., Bainbridge, D., Nichols, D. M. *How to Build a Digital Library*. Second Edition. Burlington, MA: Morgan Kaufmann Publishers, 2009; Lesk, M. *Understanding Digital Libraries*. San Francisco, CA: Morgan Kaufmann Publishers, 2005, p. 456; Kresh, D. *The Whole Digital Library Handbook* [online]. Chicago, IL: American Library Association Editions, 2007, 416 p. [Viewed 24.05.2018]. Available from: irc.tnu.edu.vn/upload/collection/brief/7805_0838909264.pdf

¹⁵ Шарон, Т., Франк, А. *Электронные библиотеки в Интернете...*, с. 131.

– *Безопасност* - Свързана е с първоначалната функция на библиотеката – съхраняването на документите (информацията) и предаването им за бъдещите поколения.

– Алекса МакГрей и Мари Галахар¹⁶ считат, че предимствата на електронната библиотека в сравнение с класическата такава са:

1. *Възможност за споделяне на информацията.* В традиционните библиотеки често издания и архивни документи са в единични бройки, което поставя ограничение за ползване от по-голям брой читатели. Дигиталните копия, споделени в мрежа, са достъпни до много хора едновременно.

2. *Нови електронни форми и формати за предоставяне на информацията.* В традиционните библиотеки почти цялата информация съществува на хартия или на други традиционни носители, които не винаги са най-подходящи за копиране, ползване и разпространение. Електронната библиотека предоставя информация в множество и различни формати.

3. *Лесно обновяване на информацията.* При книжните документи и други традиционни материали обновяването става бавно и изисква сериозно финансиране. Обновяването на дигиталните документи е лесно и бързо, като може да се посочи версията и датата на направената промяна.

4. *Достъпност навсякъде и по всяко време.* За използване на традиционните библиотеки е нужно потребителите да посетят сградата, в която се помещава библиотеката, в определен часови диапазон. Електронните библиотеки са достъпни чрез уеб и предоставят информация по всяко време и от всеки компютър, свързан с глобалната мрежа.

5. *Мощността на компютъра в помощ на търсенето на информация.* За разлика от традиционните документи, където за откриване на някакъв факт е нужно читателят да изчете всички страници, за да го открие, за компютъра търсенето отнема минимално време. Компютърните методи за търсене, макар и несъвършени, са по-бързи и точни от ръчните. Те са особено полезни за многократно повтарящи се справки.

Някои изследователи¹⁷ изразяват несъгласие с „библиотековедската” концепция за електронните библиотеки. Според тях основният „обект”, с който работят библиотеките са печатните документи и други традиционни материали с присъщите им специфики на съхраняване, обработване и използване. Електронните библиотеки, от своя страна,

¹⁶ McGray, A., Gallagher, M. Principles for Digital Library Development [online]. In: *Communications of the ACM*, 44, 2001, № 5, p. 49–54. [Viewed 15.04.2018]. Available from: https://www.researchgate.net/publication/220422376_Principles_for_Digital_Library_Development .

¹⁷ Горный, Евгений, Вигурский, Константин. Цит. съч., с. 161.

акумулират електронна информация, а тя предполага различни средства за манипулиране с нея.

1.1.2. Електронна библиотека като информационна система

Съществуват източници¹⁸, които правят интересни съпоставки на електронната библиотека с уеб. Ще ги представим накратко, защото открояват важни характеристики на електронната библиотека.

Уеб е виртуална мрежа от взаимно свързани хипертекстови документи, достъпни чрез интернет. Документите в уеб могат да съдържат текст, изображения, аудио, видео и мултимедия, а преминаването от един документ към друг се осъществява с помощта на хипервръзки. Уеб няма централизирано управление нито от гледна точка на техническата реализация, нито в областта на създаването, подбора, достъпа и използването на информация. Това го отличава от електронната библиотека, за която присъщи качества са организация и подборност с оглед на потребителските интереси.

В паралелното разглеждане се откроява още едно предимство, касаещо цифровото неравенство в обществото – като обект с краен обем електронната библиотека може да съществува върху преносим носител (като CD-ROM или DVD и др.) и да дава достъп до информация, без да е необходим интернет и скъпо мрежово оборудване¹⁹.

Обсъждането на електронната библиотека като специфична автоматизирана информационна система* се основава на събраните и анализирани в процеса на нашето изследване определения на понятието от различни литературни източници:

Digital Library Federation (1998):

Електронните библиотеки са системи, предоставящи ресурси и специализиран персонал, който подбира, структурира и предлага достъп до колекции в цифров вид с цел интерпретиране, разпространяване, съхраняване в цялостен вид и гарантиране на

¹⁸ Bainbridge, David, Nichols, David M. *Digital Libraries in Education: Specialized Training Course: Study Guide* [online]. Moscow: Unesco Institute For Information Technologies In Education, 2006, с. 14. [Прегледан 06.07.2017]. PDF, 1.1 MB. Достъпно от: <http://iite.unesco.org/pics/publications/en/files/3214563.pdf>.

¹⁹ Ibid.

* Автоматизираната информационна система е комплекс от компютърно и комуникационно оборудване, програмно осигуряване, лингвистично осигуряване, информационните ресурси и потребителите.

тяхната устойчивост, така че да са лесно достъпни и икономически използвани от определена общност или група общности²⁰.

Т. В. Ершова и Ю. Э. Хохлов (1999):

Електронната библиотека е разпределена информационна система, позволяваща надеждно съхраняване и ефективно използване на разнообразни колекции от електронни документи (текст, графика, аудио, видео и др.), достъпни в удобен за крайния потребител вид, чрез глобалната мрежа за предаване на данни²¹.

Уилям Армс (2000):

Управляема колекция от цифрова информация, достъпна чрез мрежата в съвкупност със съответстващите услуги²².

О. Голицына и Н. Максимов (2000):

Електронната библиотека е комплекс от взаимосвързани хетерогенни масиви от информация и средства, обезпечаващи достъп на ползвателите до тях във функционално и организационно еднородна среда (обикновено автоматизирана информационно-търсееща система). В общия случай компонентите на електронната библиотека образуват тристепенна структура, която включва: първична информация /пълнотекстови масиви/, вторична информация /електронни каталози или библиографски бази от данни/, метаинформация /справочно-нормативни файлове/²³.

О. В. Сьютюренко и Ю. Е. Хохлов (2000):

Електронната библиотека е система с унифициран подход към производството, съхраняването и организацията на разнообразна информация с цел търсене, анализ и достъп до нея чрез използване на глобалната компютърна мрежа²⁴.

²⁰ *Digital Library Federation* [online]. [Viewed on 19.08.2017]. Available at: <http://old.diglib.org/about/dldefinition.htm>

²¹ Ершова, Т. В., Хохлов, Ю. Е. Межведомственная программа „Российские электронные библиотеки“ : Подходы и перспективы. В: *Электронные библиотеки*, 1999, Т. 2, Вып. 2, с. 29–33.

²² Arms, William Y. *The Digital Libraries*. Cambridge, Massachusetts : MIT Press, 2000, p. 11.

²³ Голицына, О. Л., Максимов, Н. В. Технология и средства доступа к распределенным ресурсам электронных библиотек и баз данных. В: *Электронные библиотеки*, 2000, Т. 3, вып. 1, с. 16–23.

²⁴ Сьютюренко О. В., Хохлов Ю. Е. Распределенные библиотечные сети и электронные библиотеки в России. В: *Электронные библиотеки*, 2000, Т. 3, вып. 5, с. 6–10.

Digital Libraries: The Future Directions for European Research Programme (2001):

Система, която позволява на всеки достъп до цялото човешко знание по всяко време и навсякъде, по удобен многовариантен, ефективен и ефикасен начин, като преодолява пространствените, езиковите и културните препятствия чрез използване на многобройните механизми на Интернет²⁵.

А. Б. Антопольский (2002):

Електронната библиотека е информационна система, която позволява надеждно съхраняване и ефективно използване на разнообразни колекции от електронни документи (текстови, изобразителни, аудио-, видео- и др.), локализиращи в самата система или достъпни чрез телекомуникационна мрежа²⁶.

А. И. Земсков (2007):

Електронната библиотека е създадена на професионална основа тематично ориентирана (или организирана по друг начин) система за достъп до отдалечени или локални електронни ресурси, обслужваща с ресурси и специализирана помощ локални и отдалечени ползватели²⁷.

Борислав Павлов (2012):

Система, която интегрира на различни нива цифрово съдържание и метаданните за него, средства за управление на информационни ресурси и на информационни услуги, като осигурява ефективен и гъвкав достъп до знанието чрез поддръжка на разнообразни дейности, процеси, форми и формати за представяне на информационното съдържание, както и на взаимодействието с него²⁸.

Мария Нишева-Павлова (2013):

²⁵ Цит. по: **Павлов**, Радослав. *Цифрово съдържание и цифрови библиотеки за културно-историческо наследство* [онлайн]. [Прегледан 19.08.2018]. Достъпно от: mdl.cc.bas.bg/docs/Presentation_NF13CBulgaria.ppt

²⁶ **Антопольский**, А. Б. *Лингвистическое обеспечение электронных библиотек*. Москва: Информрегистр, 2002, с. 26.

²⁷ **Земсков**, А. И., **Шрайберг**, Я. Л. *Электронная информация и электронные ресурсы*. Москва: Фаир, 2007, с. 348.

²⁸ **Павлов**, Радослав. *Цифрови библиотеки с научно съдържание – реализации и проекти* [онлайн]. [Прегледан на 14.05.2017]. Достъпен от: racpl.org/wp-content/uploads/2014/02/10_R.Pavlov-DL-Science.pptx

*Електронната библиотека означава библиотека, чиито колекции се съхраняват в цифрови (дигитални, електронни) формати и са достъпни с помощта на компютър. Съдържанието на такава библиотека може да се съхранява локално или да е осигурен отдалечен достъп до него в рамките на система от компютърни мрежи*²⁹.

За работно определение в рамките на дисертационното изследване се спряхме на следното определение: **Електронната библиотека е информационна система, обикновено интернет базирана, предназначена за съхраняване и ефективно използване на разнообразни колекции от документи (текст, изображение, аудио, видео и др.), обединени в единна организационна структура и достъп.**

Определението предполага информационна система с архитектура и програмна среда, организация (или отделно лице), която извършва подбор, съхранение и т.н. и гарантира качество на информацията и услугите, политика за използване на електронната библиотека и потребители. Потребителите на електронната библиотека са както крайните потребители, така и професионални потребители („библиотекари“). Според тяхната „роля“ (потребителски профил) те могат да използват различни функции на системата. Обикновено електронната библиотека предоставя на крайните потребители услуги като достъп, търсене, навигиране, представяне и получаване на информационни обекти (документи).

Основните задачи на електронната библиотека са интеграцията на информационни обекти (документи и метаданни за тях), предоставянето на достъп, ефективно търсене и навигиране, както определя А. Антопольский³⁰.

Под „интеграция“ се разбира обединяване на разнообразна по вид информация от различни източници със запазване на нейните специфични свойства, особености и възможности за работа с нея (чрез унифициран потребителски интерфейс). При това интегрирането не се извършва непременно физически, то може да бъде и виртуално. Интеграцията се осъществява чрез въвеждане на общи средства за идентификация и описание на документите. Прилагането на общ информационен модел и протокол за взаимодействие позволява обединяването в концептуално единна система. Основното е

²⁹ Нишева-Павлова, Мария. *Създаване на хетерогенни цифрови библиотеки, основано на онтологии*: Хабилитационен труд. София, 2013, с. 19. [Прегледан 22.05.2017]. Достъпно от: https://www.fmi.uni-sofia.bg/habil_disert_trudove/habilitacionni_trudove_papka/Habilit_trud_MN.pdf

³⁰ Антопольский, А. Б. Лингвистическое обеспечение электронных библиотек. Москва: Информрегистр, 2002, с. 26.

достъпната за ползвателя информация и услуги да се възприема като единна информационна среда.

Информационният модел представя типовете същности (обекти, понятия, концепти), взаимовръзките им (отношения), ограниченията и правилата на конкретна предметна област. Той се описва чрез специален формален синтаксис, наречен език за моделиране на данни, напр. „Същност-Връзка” (Entity-Relationship, ER)³¹. Като краен резултат се постига цялостност, взаимосвързаност и интегритет.

Електронната библиотека осигурява ефективно информационно търсене основано на комплекс от информационни езици и автоматизирани методи за пълнотекстово търсене*.

Навигацията в електронната библиотека означава възможност ползвателят да търси и открива необходимата му информация с най-висока степен на пълнота и точност в цялото достъпно информационно съдържание. Този процес предполага нагледно представяне на неговата логическата структура. Създаването ѝ изисква разрешаване на задачи като структуриране и систематизиране на ресурсите по формат, теми и пр. чрез проследяване на явни и/или неявни връзки между тях.

Информационното съдържание на електронната библиотека се представя като съвкупност от самостоятелни обекти. В електронната библиотека „атомарната” единица на съхранение обикновено се нарича информационен обект*. Това е условна единица от информация, която може да бъде:

а) информационен обект/документ – текст, изображение, аудио, мултимедия документ, триизмерен обект, база от данни, колекция от отделни информационни обекти и др. Той може да има проста структура, но също и сложна вътрешна организация, като се състои от части, различни файлове и т.н.;

б) информационен обект, който е предназначен да описва ресурс (обикновено първичен информационен обект) – библиографско описание или описателни метаданни;

³¹ **Chen**, P. P. The Entity-Relationship Model – Towards a Unified View of Data. In: *ACM Transactions on database Systems* [online]. Vol. 1, No.1, March, 1976. [Viewed 19.08.2017]. Available from: <https://pdfs.semanticscholar.org/1348/1add6a6ce710af7004577ccb0ea7f92e887d.pdf>

* В семантичните електронни библиотеки се използват развити методи и средства за търсене като семантично търсене и извличане на знания от текст, изображения и потребителски анотации; многоезиково търсене; възможност за търсене в няколко системи, основано на синтактична и семантична съвместимост; намиране на отговори, базирани на логически изводи и разсъждения; адаптивно търсене и др.

* В научната литература често „информационен ресурс” и „ресурс” се употребяват като смислово еквивалентни, но тук съзнателно избягваме подобна тъждественост. „Ресурс” според „Референтния модел на електронната библиотека” е всяка същност в универсума на електронната библиотека, която може да бъде идентифицирана, напр. „колекция”, „участник”, „онтология”, „архитектурен компонент” и пр.

в) информационен обект, който пояснява, дава бележки, коментира ресурс – потребителска „анотация”.

За да стесним смисъла на „информационен обект” в настоящото изследване ще използваме термина „документ”, когато става дума за първични информационни обекти, независимо от техния тип медия и вид. За метаинформационните обекти, тоест структурираната информация, необходима за управление, представяне, идентифициране, достъп и съхранение на документите, ще използваме понятията „библиографско описание” или „метаданни”.

Документите в електронната библиотека трябва да бъдат представени и индексирани като логически единици чрез библиографско описание, което описва техните характеристики и да бъде предствено тяхното съдържание чрез специални информационни езици. Библиографското описание на документите (метаданните) се съхраняват в хранилище за метаданни – част от електронната библиотека. Хранилището включва функции за управление на метаописанията и позволява изпълнението на заявки за търсене по елементите на библиографското описание, включително по индексите за съдържанието на документите³². Най-използваната форма на организация на документите в електронната библиотека е колекцията. Тя представлява систематизирана съвкупност от документи, обединени по някакъв критерий за принадлежност: вид на документите, съдържание, предназначение, източник, начин на достъп и др. Колекциите могат да бъдат тематични, хронологични, да са формирани по тип на информацията (текстови, картографски, нотни и т.н.) или да бъдат библиографски. Колекцията предполага единни механизми, начини и средства за съхранение, индексирание, представяне, търсене и преглеждане. Електронната колекция има определена самостоятелност, освен това може да има своя вътрешна структура – да съдържа подколекции или да влиза в по-големи обединения на няколко колекции. Всяка колекция може да има собствен профил на метаданни и свои лингвистични средства.

Обедняването на колекциите задължително се основава на комуникативен формат на метаданни, напр. Dublin Core. Възможно е към базовите елементи в схемата да се добавят допълнителни уточняващи „квалификатори” за отделните колекции.

³² *Дигитални библиотеки с мултимедийно съдържание и приложения в българското културно наследство*: [онлайн]. Изследователска разработка по договор ИД 8/21.07.2005 с Агенция „Развитие на съобщенията и на информационните и комуникационните технологии” I етап, с. 39–40. [Прегледан на 13.12.2016]. Достъпно от: http://mdl.cc.bas.bg/Digital_libraries_with_multimedia_content_and_applications_in_Bulgarian_cultural_heritage.pdf

Търсенето в рамките на една колекция се провежда в съответствие с профила на метаданните за тази колекция, с използване на контролирани списъци и нормативни файлове, което осигурява значително по-висока релевантност на търсенето. Търсенето в цялата електронна библиотека се осъществява по общите елементи на метаданните, специфичните елементи от метаданни за отделните колекции могат да бъдат игнорирани.

Основните функции, присъщи за електронните библиотеки още в ранния им етап на развитие, са:

- добавяне, съхраняване, редактиране и изтриване на документи и метаданни;
- достъп и показване на документ;
- търсене и навигиране;
- регистрация на потребителите³³.

С развитието на електронната библиотека те се обогатяват и развиват:

– административни услуги – управление на достъпа, на потребителски профили, на дейностите и др.;

- защита на информационното съдържание;
- експорт на информационни обекти;
- обновяване и допълване на функции за агрегиране на информационни обекти от различни електронни библиотеки и хранилища с еднородно или разнородно съдържание и др.³⁴

Семантичните електронни библиотеки, които могат да бъдат назовани „библиотеки от последно поколение”, предлагат на потребителите функции и услуги, основани на технологии на Уеб 2.0 и на семантичния уеб.

Ролята и функциите на електронната библиотека предопределят някои важни изисквания при нейното изграждане, особено ако е с научно или културно-историческо съдържание. Те са формулирани от български изследователи³⁵ така:

- Електронната библиотека трябва да бъде конструирана като високотехнологична и да покрива широка област на приложение, за да задоволява интересите на различните

³³ Павлов, Радослав. Цифрови библиотеки с културно-историческо и научно съдържание [онлайн]. [Прегледан 14.05.2017]. Достъпно от: mdl.cc.bas.bg/radko/rp/Pavlov-lecture_final.ppt

³⁴ Павлов, Радослав. Цифрови библиотеки с културно-историческо ...

³⁵ Дигитални библиотеки с мултимедийно съдържание и приложения в българското културно наследство: Изследователска разработка по договор ИД 8/21.07.2005 с Агенция „Развитие на съобщенията и на информационните и комуникационните технологии”. I етап [онлайн]. с. 39–40. [Прегледан 13.12.2016]. Достъпно от:

http://mdl.cc.bas.bg/Digital_libraries_with_multimedia_content_and_applications_in_Bulgarian_cultural_heritage.pdf

потребители. Тя трябва да предвижда продължаващо присъединяване на нови медийни обекти и нови процедури за описание на тяхната семантика и семантичното им аотиране;

- Електронната библиотека трябва да поддържа средства, които да предоставят семантична контекстно-базирана информация, която би позволила всички материали, свързани с определена тематика, да бъдат откривани и групирани в подходяща за разглеждане форма;

- При създаването на електронна библиотека е важно учените и изследователите от различните сфери на човешкото знание да обединят усилията си и да намерят общ език с компютърните специалисти, създаващи базата от данни;

- Електронната библиотека трябва да бъде устойчива, което предполага, че е необходимо нейният архитектурен модел да бъде променяем, разширяем и да работи максимално функционално;

- Препоръчително е да се дефинира унифицирана терминология, която максимално точно да определя използваните понятия в системата. По този начин ще се гарантира цялостно разбиране на информацията от широк кръг потребители.

- За осигуряването на широка употреба на електронните библиотеки в световен мащаб е необходимо те да са разработени съгласно утвърдени стандарти и спецификации за представяне на информацията и знанието;

- Електронната библиотека трябва да съблюдава законови рамки по отношение на произход, създател, авторски права, поверителност.

1.2. Типология на електронните библиотеки

Общоприета класификация на електронните библиотеки, която да отчита разнообразните особености и характеристики на различните електронни библиотеки, не съществува. Без да претендираме за изчерпателност, на база на най-съществените признаци предлагаме следната типологизация, при която е възможно и припокриване между отделни типове електронни библиотеки:

а) Според метода на формиране на фонда електронните библиотеки могат да бъдат три типа:

- генериращи съдържание – документите и метаданните се създават от организацията собственик на електронната библиотека (напр. *eLibrary** – електронната

* Световната банка провежда и публикува проучвания и научни изследвания от широк кръг глобални и регионални въпроси като социално развитие, социална защита и труд, намаляване на бедността и др., които предоставя за достъп чрез своята електронна библиотека (<https://elibrary.worldbank.org/>).

библиотека на Световната банка). В общия случай това са електронни библиотеки на научни, изследователски и образователни организации;

– агрегиращи съдържание – документите и метаданните се събират и обединяват в единно информационно съдържание (напр. Европейската електронна библиотека *Европеана***);

– смесени – използват двата подхода за съставяне и поддържане на информационното съдържание.

б) Според видовете документи, които се включват в електронната библиотека можем да разграничим: електронни библиотеки, които съдържат само един вид документи или един тип информационни обекти (например *Библиотека дисертаций и авторефератов России****);

– електронни библиотеки, които съдържат разнообразни типове и видове документи или информационните обекти: според знаковата природа на документите (текстови, видео, мултимедия и т.н.); според по начина на създаване на електронните документи (дигитализирани и изначално дигитални), например *Digital library for Earth System Education***** предлага над 60 категории електронни ресурси (като проекти, игри, карти, тестове, учебни курсове и т.н.).

в) В организационен аспект електронните библиотеки като информационни системи могат да бъдат:

– самостоятелни – представляват отделни системи за електронни информационни ресурси;

– част от по-голям ресурс или информационна система с по-общо предназначение (напр. тематичен портал, система за вземане на решения, портал за електронно обучение);

г) Според целевото си предназначение електронните библиотеки могат да бъдат*:

– научни – предназначени за изследване на определена научна област, дисциплина, тема от учени, изследователи и специалисти (напр. електронната библиотека на

**Европейската електронна библиотека *Europeana* функционира като уеб портал за цифрова културна продукция, който свързва цифрови сбирки на културното наследство на държавите от ЕС.

*** <http://www.dslib.net/>

**** <http://www.dlese.org/lib/index.html>

* Делението е съвсем условно, защото отделните колекции в една електронна библиотека могат да имат различна потребителска насоченост.

Американската математическа асоциация - Mathematical Sciences Digital Library, /MathDL/**);

– учебни – предназначени за поддържане и подпомагане на образованието и обучението (напр. електронните библиотеки на университети; електронни библиотеки, които включват информационни обекти за различни степени на образование като Digital library for Earth System Education (DLESE))***.

– справочни – електронни библиотеки подобни на универсални енциклопедии, предназначени за получаване на необходима кратка информация от различни отрасли на знанието;

– научно-популярни – предназначени за масовия потребител за самообразование и личностно развитие (напр. Виртуална духовна библиотека „Спирала на еволюцията“****);

– мемориални – създадени с цел да представят лице или събитие (напр. Brill's Digital Library of World War I*****);

– служебни (или корпоративни) - обслужват информационното осигуряване на конкретна организация с нормативни документи, бизнес информация и т.н.;

– развлекателни – съдържат игри или други интерактивни информационни ресурси за развлечение.

д) В зависимост от създателя си електронните библиотеки се разделят на:

– библиотеки, създадени от органи на властта, министерства, ведомства, национални институции (напр. Дигитална Библиотека на Национален статистически институт*);

– библиотеки, създавани от институциите на паметта - библиотеки, архиви, музеи и др. (напр. Дигитална библиотека на Национална библиотека „Св. св. Кирил и Методий“**, Дигитален архив на Национална агенция „Архиви“***, електронна библиотека на музея в Чарлстън, Южна Каролина, САЩ - Digital Library Charleston Museum****);

** <https://www.maa.org/about-maa/maa-history/mathdl>

*** Библиотеката се поддържа от съобщност от преподаватели, студенти и учени, които подпомагат образованието във всички степени (Elementary School, Middle School, High School)

<https://www.dlese.org/lib/index.html>

**** <http://www.spiralata.net/>

***** <https://brill.com/view/db/lwso>

* <http://statlib.nsi.bg:8181/>

** http://www.nationallibrary.bg/wp/?page_id=1337&lang=bg

*** <http://www.archives.government.bg/>

**** <https://www.charlestonmuseum.org/research/digital-library/>

– библиотеки, създадени от средствата за масова информация, от обществени структури, от научни и образователни институции, от издателства, от комерсиални организации, от отделни лица, от граждански обединения и пр.

е) Според съдържанието на информационния фонд електронните библиотеки могат да бъдат:

- универсални;
- по отрасъл, направление, научна област;
- по определена тема.

ж) Електронните библиотеки могат да се различават и по начина на достъп:

- електронни библиотеки със свободен достъп;
- електронни библиотеки с достъп след потребителска регистрация;
- електронни библиотеки с локален достъп (в локална мрежа или на физически носител).

Предложената типология далеч не обхваща всички признаци, като се има предвид големия брой учреждения, организации, органи и дори отделни лица, които създават свои разнообразни колекции от електронни документи и ги предоставят за достъп. Електронните библиотеки възникват спонтанно в интернет средата и не изискват, както е при традиционните библиотеки, регистрация като отреден тип и вид библиотека в съответствие с нормативните документи. Освен това е много трудно тяхното идентифициране. От една страна, „електронна библиотека” може да бъде назован всеки сайт или раздел в сайта, съдържащ набор от дигитални документи. От друга страна, не всеки създател на електронна библиотека счита за необходимо да постави в заглавието думата „библиотека”.

1.3. Концептуални модели на електронна библиотека

При изграждането на устойчива електронна библиотека основни цели са:

а) осигуряването на оперативна съвместимост с други информационни системи (в частност други електронни библиотеки);

б) възможност за многократно използване на цифровото съдържание в друг контекст и област, с други дигитални колекции.

Съвременните технологични решения за създаване на електронни библиотеки трябва да се базират на утвърдени добри практики, стандарти и спецификации. Осъществяването на успешни реализации на електронни библиотеки се верифицира чрез

съобразяването със специално създадения референтен модел* на електронна библиотека по проект DELOS „A Network of Excellence on Digital Libraries”, DL.org (Coordination Action on Digital Library Interoperability, Best Practices and Modelling Foundations)³⁶ и др.п.

1.3.1. Референтен модел на електронна библиотека DELOS DLRM

Референтен модел на електронната библиотека DELOS Digital Library Reference Model (DLRM). Моделът предлага обща концепция при експертно съгласие за споделено разбиране за основните понятия, структурата и функциите на електронната библиотека³⁷. Създаването на рамка, базов модел, по който могат да се съпоставят електронните библиотеки или проекти за електронни библиотеки е изходна точка за постигане на способност за комуникация (взаимодействие) между електронните библиотеки³⁸. Взаимодействието може да има различни форми – напр. прехвърляне, обменяне, трансформиране или интегриране на информация³⁹.

Референтният модел на електронната библиотека е резултат от дейността на няколко европейски изследователски групи по проект, подкрепен от Европейската комисия, и завършен през 2007 г. Моделът е публикуван в редактирана и разширена версия през 2011 г. в обширен документ от 4 обособени части, значими сами по себе си: „Манифест на електронната библиотека”, „Рамка на „Референтния модел на електронната библиотека”, „Референтен модел на електронната библиотека” и „Списък с критерии за съответствия на „Референтния модел на електронната библиотека”.

Според референтния модел електронната библиотека се разглежда като трислойна „структура” (фиг. 1), която се състои от:

Електронна библиотека (Digital Library, DL) – организация (включително и виртуална), която събира, обработва, управлява и съхранява дългосрочно цифрово съдържание и предлага на различни типове потребители специализирани функционални

* Референтен модел (концептуален или еталонен модел) най-общо представлява абстрактно описание, рамка на понятията и отношенията между тях в някаква предметна област. Целта на всеки един референтен модел е да изрази еднозначно смисъла на термините и понятията (концептите), използвани от експертите в дадена област и да отрази подходящите връзки между тях. Референтният модел е инструмент за съгласувано разбиране и прилагане на решения. Не е пряко свързан с никакви стандарти, технологии или други конкретни подробности за изпълнението, но има за цел да осигури обща семантика, която може да бъде използвана недвусмислено във и между различните реализации.

³⁶ Желев, Янислав, Тошков, Ангел, Николова, Евгения, Желева, Мария. Развитие на икономика, базирана на знания и иновации В: Бизнес посоки [онлайн]. 2015, 1, с. 88–99. [Прегледан 06.07.2017]. Достъпно от: http://research.bfu.bg:8080/jspui/bitstream/123456789/617/1/BFU_BP_2015_1_Zhelev_Toshkov_Nikolova_Zheleva.pdf

³⁷ Candela, L., Nardi, A. eds. Op. cit. p. 56.

³⁸ Пак там.

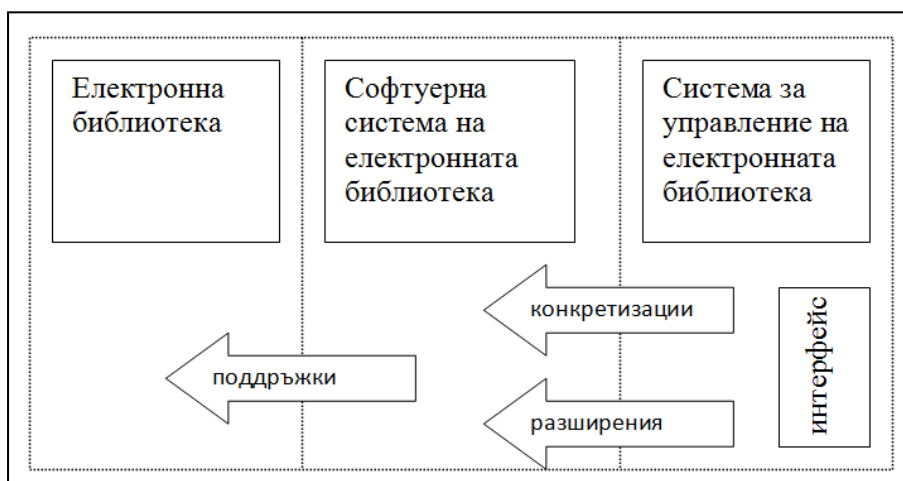
³⁹ Нишева-Павлова, Мария. Създаване на хетерогенни цифрови библиотеки... [онлайн]. с. 24.

възможности за достъп и използване на това съдържание в съответствие с добре дефинирани критерии за *качество* и ясно определени *политики*.

Система на електронната библиотека (Digital Library System, DLS) – програмна среда, основана на определена архитектура, която реализира функционалните възможности изисквани от електронна библиотека. Крайните потребители общуват с електронната библиотека чрез системата, в чиято среда се създава конкретна електронна библиотека (например популярните DSpace, EPrints и др.).

Система за управление на електронна библиотека (Digital Library Management System, DLMS) – платформа за администриране на програмната система на електронната библиотека с възможности за интегриране на допълнителен софтуер, която позволява разширяване и обогатяване на поддържаните функционални възможности (напр. OpenDLib Digital Library Management System). Със системата за управление общуват проектантите, разработчиците и администраторите на електронната библиотека.

Фигура 1. Референтна архитектура на електронна библиотека⁴⁰



Основните изграждащи елементи на електронната библиотека са:

Организация (Organization) – понятието се свързва с целите и мисията на електронната библиотека и се разграничава по смисъл от организацията учреждение, която създава и стимулира развитието на електронната библиотека, макар и да има известно припокриване. „Организация” е супер-концепт, който обхваща останалите основни концепти в електронната библиотека, които се свеждат до понятието „ресурс”.

⁴⁰ Пак там, с. 23.

Архитектура (Architecture) – отнася се до програмната система на електронната библиотека. Зависи от функциите и съдържанието, предлагани от електронната библиотека и се основава на технически и програмни средства.

Политика (Policy) – представлява съвкупност от условия и правила, които управляват отношенията между електронната библиотека и нейните потребители. Примери за политики са управлението на права, поверителността, таксите за използване и т. н.

Съдържание (Content) – обхваща данни и информация, които електронната библиотека обработва и предоставя на потребителите. Понятието включва разнообразни информационни обекти, включително метаданни.

Функционалност (Functionality) – включва функционалните възможности, които предоставя електронната библиотека на различните категории потребители.

Потребители (User) – включва различните участници, които взаимодействат с електронната библиотека. Те могат да ползват информация или да генерират информация (независимо дали са хора или машини). Тук се включват правата в системата на електронната библиотека, профилите на потребителите с персонализиращи характеристики и др.

Качество (Quality) – понятие, което обединява параметри, които могат да бъдат използвани за характеризирание и оценка на съдържанието, функционалните възможности, информационни обекти и др.

За да се опише цялостното функциониране на една електронна библиотека, задължително трябва да се отрази участието на различните категории потребители:

– „Крайни потребители” (DL End-users): „създатели на съдържание” (Content Creators), „потребители на съдържание” (Content Consumers) и „библиотекари” (Digital Librarians) – професионални потребители, които подбират и организират съдържанието.

– Потребителите, които управляват програмната система на електронната библиотека (DL Managers) се делят на „проектанти на програмата” (DL Designers) и „системни администратори” (DL System Administrators).

– Потребителите, които взаимодействат със системата за управление на електронната библиотека са „разработчици на приложения” (DL Software Developers).

Общо всички концепти, които дефинира моделът са 226. Те са представени с техните свойства и във връзки с други понятия, разпределени по групи или области (Domain) и са онагледени чрез т.нар. концептуални карти. „Областите” се изграждат около основните концепти: Електронна библиотека, Съдържание, Потребители, Функции, Политика,

Архитектура, Качество. Всяка „област“ отговаря за определен аспект, свързан с функционирането на електронната библиотека, затова ще разгледаме представляващите интерес за настоящето изследване – съдържание, потребители и функционални възможности.

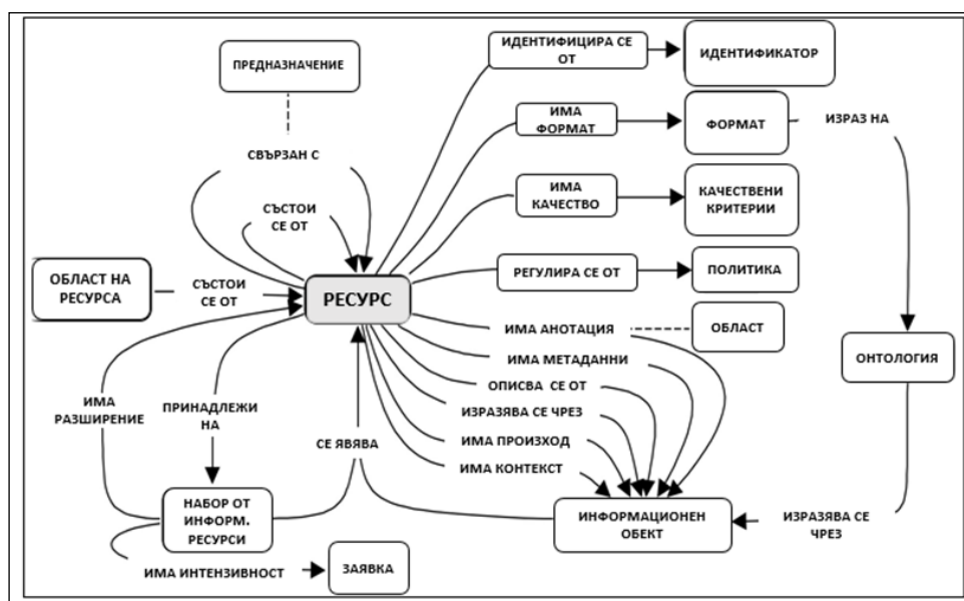
Областта на ресурса (Resource Domain) е най-широкообхватна, тъй като всички концепти се определят като „ресурс“. Според дефиницията в модела „ресурс“ (Resource) е всяка единица в универсума на електронната библиотека, която може да бъде идентифицирана*. „Ресурс“ е обща, абстрактна същност, която може да бъде представена чрез конкретизиране на една от нейните специализации, например „колекция“, „участник“, „онтология“ и пр.

Всеки „ресурс“ има следните свойства (характеристики):

- идентифицира се чрез идентификатор (<identifiedBy>);
- организиран е в съответствие с формат (<hasFormat>). Форматът може да бъде определен чрез онтология, която гарантира еднозначна интерпретация. „Ресурсът“ може да бъде сложно структуриран, да се състои от по-малки ресурси (<hasPart>) и да бъде свързан с други ресурси (<associatedWith>);
- изразява се чрез информационен обект (<expressedBy>);
- всеки ресурс може да се опише и допълни с метаданни (<hasMetadata>), контекст (<hasContext>) и произход (<hasProvenance>);
- може да се характеризира чрез качествени критерии (<hasQuality>);
- регулира се чрез политики (<regulatedBy>), които управляват неговия жизнен цикъл в електронната библиотека;
- може да принадлежи към набор от ресурси (<belongTo>), който също се разглежда като ресурс, например „колекция“ в областта „съдържание“ или „група“ в област „потребители“ (фигура 2).

* „Ресурс“ се идентифицира чрез идентификатори като Uniform Resource Identifiers (URIs), Internationalized Resource Identifiers (IRIs), Archival Resource Keys (ARKs), Digital Object Identifier (DOI) или други.

Фигура 2. Концептуална карта на областта на ресурса*.



Областта на съдържанието (Content Domain) – обхваща концепти, които се отнасят до информацията в електронната библиотека. Тук централно място заема понятието „*информационен обект*” (Information Object). То се разглежда като всяка единица информация, управлявана от електронната библиотека като текст, изображение, звук, мултимедиен документ, база от данни и др. (Фиг. 3).

Информационните обекти се специализират според функцията, която изпълняват по отношение на други информационни обекти:

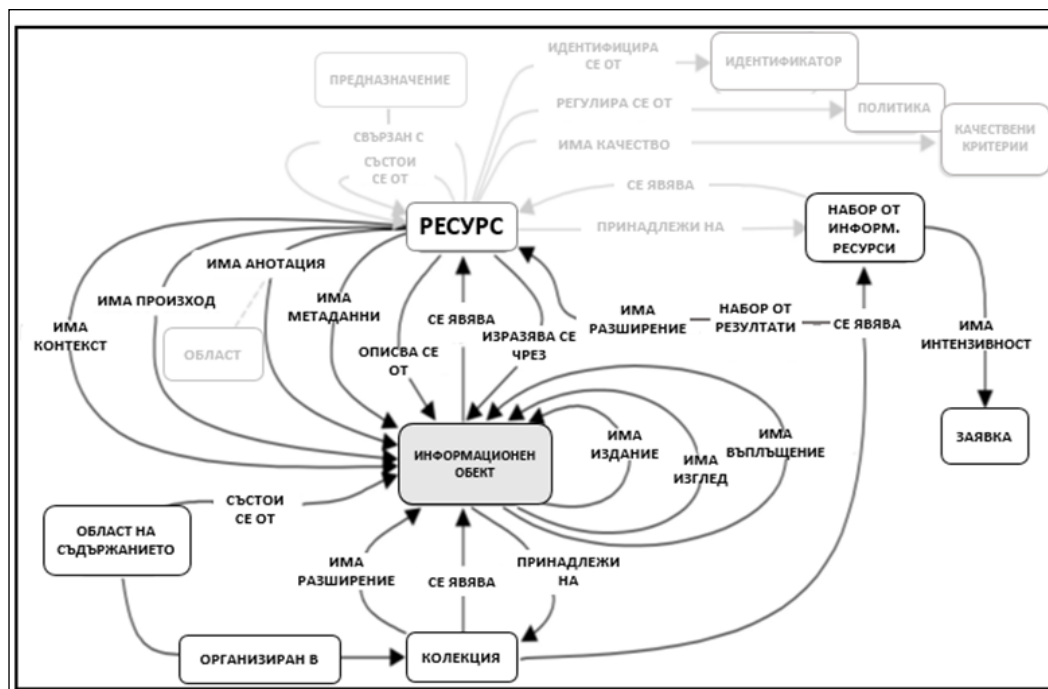
- *първичен информационен обект* – информационен обект, който се използва самостоятелно, напр. текстови документ или изображение;
- *метаданни* – информационен обект, чието доминиращо значение е да представи информация за ресурс (обикновено за първичен информационен обект);
- *анотация* – информационен обект, чиято главна задача е да дава кратка характеристика на „ресурс” в частност, напр. на документ. Примерите за анотация включват бележки, структурирани коментари и връзки. Анотацията подпомага интерпретацията на ресурса.

Моделът разглежда и други типологизиращи критерии за „информационните обекти” като:

*Източник: Candela, L., Nardi, A. eds. Op. cit. p. 39.

- *ниво на абстракция*: произведение, проявление, въплъщение, единица, взаимствани от концептуалния модел FRBR*;
- *по начин на представяне на информацията (кодиране)*:
 - ⇒ информационни обекти, в които информацията/знанията са представени в обичайна форма като документ (в най-широк смисъл тук се включват и звуковото и графичното представяне);
 - ⇒ информационни обекти, в които информацията/знанията са кодирани във вид на формални структури, напр. като таблици на база от данни;
- според това дали съдържат данни или информация и знания:
 - ⇒ информационни обекти, съдържащи необработени данни (данни и потоци от данни, събрани от каквито и да било инструменти);
 - ⇒ информационни обекти, съдържащи обработени или генерирани от човек или система данни, определяни като информация или знания;
- според начина на цифровото представяне:
 - ⇒ информационни обекти създадени в цифрова форма,
 - ⇒ информационни обекти в цифрова форма, създадени в следствие на дигитализация на аналогови документи,
 - ⇒ информационни обекти, които могат да бъдат представени в електронна библиотека само чрез метаданни.

Фигура 3. Концептуална карта на областта на съдържанието*



* за концептуалния модел FRBR виж, т. 1.3.2.2.

*Източник: Candela, L., Nardi, A., eds. Op. cit., p. 40.

„Информационният обект” като една от конкретизациите на „ресурс” споделя всички присъщи му свойства (фигури 2 и 3). В частност „информационният обект” трябва да принадлежи на поне една колекция (<belongsTo>). Може да има различни издания (<hasEdition>), различни изгледи (<hasView>) и различни възплъщения (<hasManifestation>), които също са информационни обекти.

„Колекцията” (Collection) е набор от информационни ресурси (Resource Set), който от своя страна също представлява „информационен обект”, следователно и „ресурс”, и има същите характеристики, напр. описва се с метаданни, може да бъде анотиран и др. Моделът разглежда „колекцията” като класически библиотечен механизъм за организиране на информационните обекти и за представяне на фонда. Те дават на потребителите възможност за достъп до обособени части от цялото, напр. по тематичен принцип. „Колекциите” могат да бъдат създавани от „библиотекари” или от упълномощени „потребители на съдържание”, които да реализират собствените си виждания за организиране на интересующите ги информационни обекти.

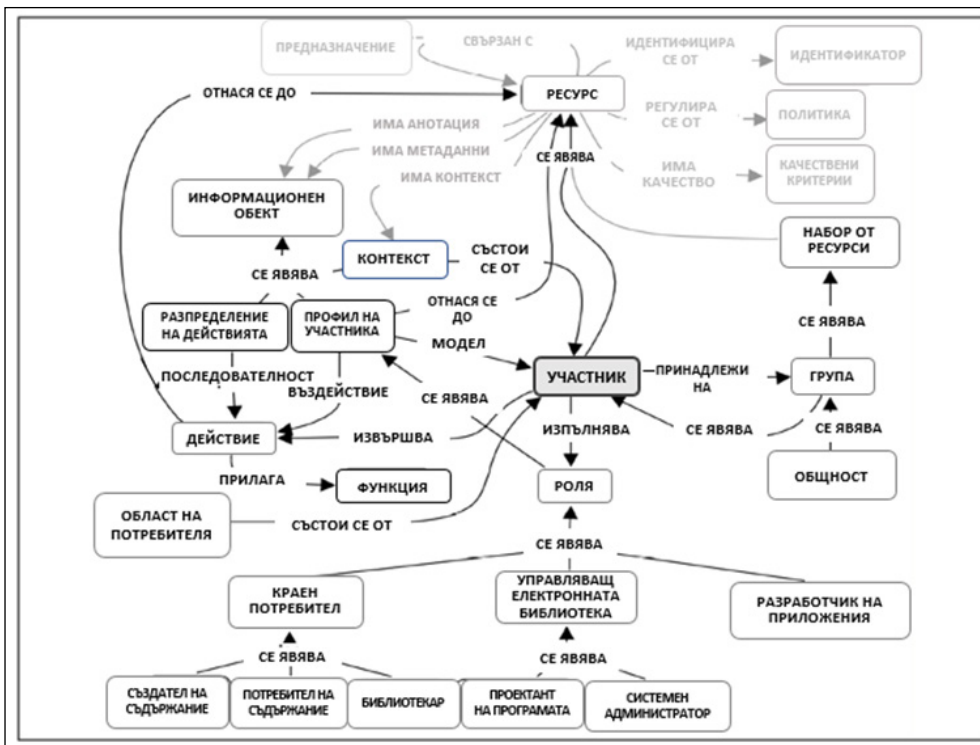
Областта на потребителя (User Domain) включва всички същности, които взаимодействат с всяка от системите на електронната библиотека: хора и неодушевени агенти, напр. програми, машини, инструменти или дори друга електронна библиотека. Главната същност в тази област е „участник” (Actor) -лице или машина. По силата на връзката (<IsA>) с „ресурс” той се характеризира със своите свойства: притежава уникален идентификатор, може да бъде описан с метаданни и др. Освен това „участник” има специфични характеристики като:

– определена „роля” (Role) като „потребител на съдържание”, „создател на съдържание”, „библиотекар”, „проектант на програмата”, „системен администратор”, „разработчик на приложения”. „Ролята” предполага предварително дефинирани функции (задачи), които един „участник” има право да изпълнява в рамките на електронната библиотека;

– „профил”(Profile) – всеки „участник” притежава поне един профил. Възможно е да притежава един профил, когато е в ролята на „создател на съдържание” и друг профил - като „потребител на съдържание”;

– „принадлежност” - „участникът” може да бъде част от група (Group) или общност (Community) (фиг. 4).

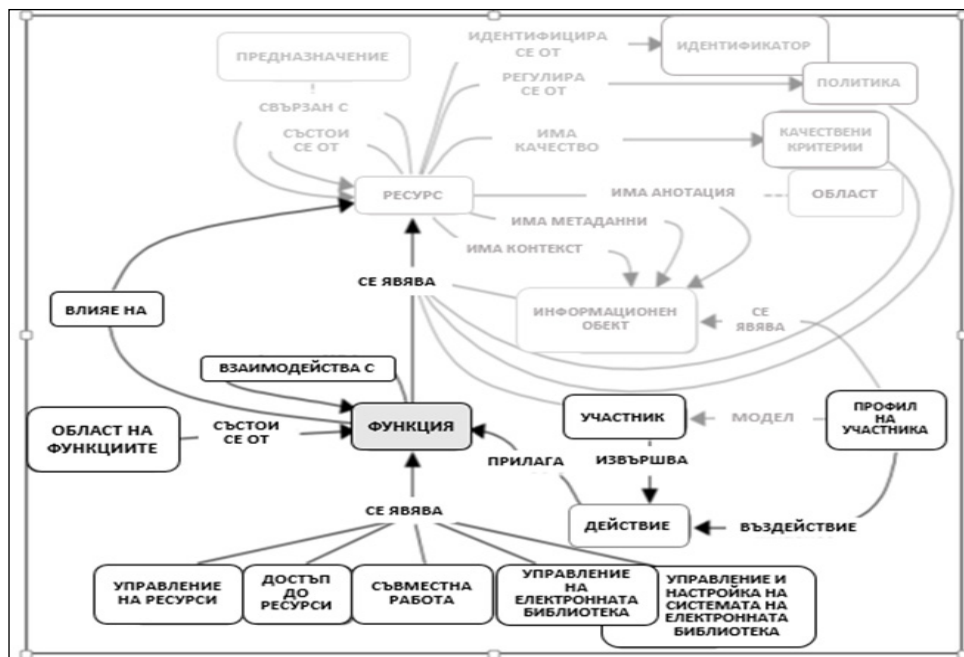
Фигура 4. Концептуална карта на областта на потребителя*



Областта на функциите (Functionality Domain) е най-отворената част в модела, тъй като представя различните аспекти, свързани с моделирането на услугите, предоставяни от електронната библиотека (фигура 5). В тази област в центъра е поставено понятието „функция”, което се разглежда като специфична задача, свързана с работата с ресурс или набор от ресурси в резултат от действие на отделен „участник”. Поради широкия обхват на „функция” не е възможно да се изброят и предвидят всички задачи и действия, които могат да бъдат включени в електронна библиотека. Отделните „слоеве” („електронна библиотека”, „система на електронната библиотека”, „система за управление на електронната библиотека”) имат собствен набор от функционални възможности, в зависимост от целите и „участниците”, които взаимодействат с тях.

*Източник: Ibid., p. 42.

Фигура 5. Концептуална карта на областта на функциите*



„Функциите” в модела са представени в пет общи групи (класове). Първите три от тях са свързани с крайните потребители (DL End user):

- *Управление на ресурси* (Manage Resource) – предвид широкия обхват на понятието „ресурс” този клас функции се разглежда в подкатегории. Като най-съществени сред тях за крайните потребители можем да определим: „Управление на информационни обекти” (Manage Information Object) – включва задачи за разпространение, обработка и др. и „Управление на участниците” (Manage Actor) – обхваща действия, свързани с регистрацията и профила на участниците;

- *Достъп до ресурси* (Access Resource) – включва всички дейности, свързани с търсене, намиране, навигиране, представяне и получаване на ресурси.

- *Съвместна работа* (Collaborate) – дейности, свързани с възможността „участници” да работят заедно, за да постигнат обща цел.

- Другите класове функции са свързани с ролята на „професионалните” потребители на електронната библиотека:

- *Управление на електронната библиотека* (Manage DL) - включва голямо разнообразие от задачи, свързани с работата на електронната библиотека, напр.

*Източник: Ibid.

„импортиране на колекция” или „управление на група”. Тези задачи се свързват преди всичко с ролята на „библиотекарите” (Digital Librarians).

– *Управление и настройка на системата на електронната библиотека (Manage & Configure DLS)* – обхваща задачи за настройване, конфигуриране, актуализиране и др. Тази група от действия и задачи се свързва с „участници”, взаимодействащи със „системата на електронната библиотека” – „проектанти” и „системни администратори”.

„Функция” също се явява „ресурс” и следователно споделя всички негови свойства.

DELOS DLRM описва рамка на електронна библиотека, която на следващ етап би гарантирала възможност за оперативна съвместимост. Моделът служи за разработване на конкретни структури и архитектури за реализация чрез програмни средства, както и за оценка на съвместимост на проекти или съществуващи електронни библиотеки.

1.3.2. Информационни модели на електронна библиотека

Ако DELOS DLRM описва съвместим модел на електронна библиотека, то група стандарти или утвърдени спецификации се създават с цел да моделират (концептуализират) информацията за определена предметна област. Като най-съществени за изграждането на електронна библиотека ще бъдат разгледани моделите за международен обмен на информация в областта на културното наследство CIDOC CRM, за функционалност на библиографските данни фамилията модели FRBR и IFLA LRM, и за научноизследователската сфера – CERIF.

1.3.2.1 CIDOC CRM

Концептуалният референтен модел CIDOC Conceptual Reference Model, (CIDOC CRM)⁴¹ е разработен от Международния комитет по документация на Международния съвет за музеите (The International Committee for Documentation of the International Council of Museums, ICOM-CIDOC) в резултат над десетгодишната дейност на работните групи „Documentation Standards Working Group” и „CRM SIG”. Представява формална онтология от най-високо ниво, предназначена да улесни интеграцията и обмена на разнородна

⁴¹ **Boeuf**, Patrick Le, **Doerr**, Martin, **Ore**, Christian Emil, **Stead**, Stephen, eds. *Definition of the CIDOC Conceptual Reference Model Version Version 6.2.3 May 2018* [online]. [Viewed 24.05.2018]. PDF. Available from: <http://www.cidoc-crm.org/Version/version-6.2.3-0> .

информация за културното наследство. „Информационната интеграция е фокусът на CRM”⁴².

За разлика от повечето онтологии, които представят експертната терминология в дадена предметна област, CRM е базова онтология (Foundational Ontology), която има за цел семантична интеграция на хетерогенни структури от данни. Както твърди един от авторите - Мартин Доер, моделът на същностите и връзките между тях (онтология) е много по-съществен за интеграцията на знанието, отколкото дефинирането на терминологията⁴³.

Експертите създават модел, съвместим със стандартите на семантичната мрежа (езика RDF) и функционалния подход към библиографската информация (FRBR). Той е разработен на езика за представяне на знания TELOS и е достъпен в RDFS и други формати. В края на 2006 г. моделът е приет от Технически комитет по информация и документация (TC46) на Международна организация по стандартизация (International Organization for Standardization) като международен стандарт (ISO 21127:2006 A reference ontology for the interchange of cultural heritage information*).

CIDOC CRM трябва да бъде общ „език” между експертите в областта на културното наследство и разработчиците на електронни библиотеки с „културна продукция” като формулира изисквания към информационните системи, служи като ръководство за концептуално моделиране и дава условия за ефективна семантична съвместимост.

Приложението на CIDOC CRM е в няколко направления:

- служи за концептуално моделиране на предметната област – културно наследство;
- базисна онтология за разработване на приложни онтологии;
- схема (онтология), която гарантира интегриране на информация и знания и обмен между комуникиращи си системи⁴⁴.

⁴² **Doerr**, Martin. The CIDOC CRM, an Ontological Approach to Schema Heterogeneity. In: *Semantic Interoperability and Integration. Dagstuhl Seminar Proceedings 04391* [online]. Dagstuhl: Schloss Dagstuhl, 2005. [Viewed 24.05.2018]. PDF, 0.2 MB. Available from: <https://pdfs.semanticscholar.org/c40f/1da72c8c952437135843a3947f56533c334c.pdf>.

⁴³ Ibid., с. 2.

* От декември 2014 г. съществува нова версия на стандарта (базирана на версия 5.0.4 на CIDOC CRM) ISO 21127:2014.

⁴⁴ **Guarino**, N. Formal Ontology and Information Systems. In: *Formal Ontology in Information Systems: Proceedings of the 1st International Conference* [online]. Trento, Italy, June 6–8 1998. Amsterdam: IOS Press. p. 3–15 [Viewed] 24.05.2018]. PDF, 58.05KB. Available from: <https://klevas.mif.vu.lt/~donatas/Vadovavimas/Temos/OntologiskaiTeisingasKonceptinisModeliavimas/papildoma/Guarino98-Formal%20Ontology%20and%20Information%20Systems.pdf>

CIDOC CRM дава определения и формална структура за описване на обекти* (entity) и свойства (property) – т.е на бинарни отношения. Версия 6.2.3 (май 2018) на модела⁴⁵ представя 99 обекта и 188 свойства. Обектите са представени в йерархичен порядък на няколко нива чрез отношението „се явява” („Is A”). Всички те са подчинени на E1 Обект (CRM Entity) – клас от най-високо ниво, който съдържа няколко суперкласове: E2 Временен обект (Temporal Entity), E52 Период (Time-Span), E53 Място (Place), E54 Размер (Dimension), E77 Постоянен обект (Persistent Item). Изключение прави E59 Стойност (Primitive Value), който влиза в нов йерархичен ред. Той съдържа данни като цифри, букви, знаци, символи и др., кодирани чрез някаква система. Неговите суперкласове са E61 Дата (Time Primitive) и E62 Низ от знаци (String) (виж табл.3).

Таблица 1. Йерархия на класовете същности (Class Hierarchy) в CIDOC CRM

E1	Обект	
E2	-	Временен обект
E77	-	Постоянен обект
E52	-	Период
E53	-	Място
E54	-	Размер
E59	Стойност	
E61	-	Дата
E62	-	Низ от знаци

Представените в модела обекти и свойства условно могат да бъдат разделени в три групи. Първата включва такива, които обхващат общи понятия от обкръжаващия ни свят. Втората група обединява обекти и свойства, свързани с управлението – придобиване и регистрация на единици за съхраняване, права и др. Към третата група се отнасят обекти и свойства, използвани за организация на самата онтология, напр. E32 Authority Document и P71 Lists (is listed in).

* *entity* (англ. ез.) – обект или същност. В това изследване е предпочетен терминът „обект”, тъй като вече е навлязъл в професионалната терминология на български език чрез преводите на FRBR (Функционални изисквания към библиографските записи) и FRAD (Функционални изисквания към нормативните данни). В някои случаи се използва като синоним „същност”, за да се избегне повторение.

⁴⁵ **Boeuf**, Patrick Le, **Doerr**, Martin, **Ore**, Christian Emil, **Stead**, Stephen, eds. *Definition of the CIDOC...*

Моделът може да бъде разширяван чрез добавяне на необходимите за конкретно изпълнение на електронна библиотека същности.

Ако DELOS DLRM разглежда основните аспекти на електронна библиотека, а CIDOC представя в концептуален обектно-ориентиран модел информацията за културно наследство, то разглежданите по-долу модели са ориентирани към ползвателите на библиографска информация.

1.3.2.2. FRBR

Концептуален модел „Функционални изисквания към библиографските записи” (Functional Requirements for Bibliographic Records, FRBR)⁴⁶. Библиографската информация има разнообразни приложения и широк кръг ползватели в съвременното информационно пространство. Този факт провокира експертите на Международната федерация на библиотечните асоциации и институции (International Federation of Library Associations and Institutions, IFLA) да проведе изследване (1991–1997 г.) относно какво ползвателите очакват да намерят информация в библиографския запис и как използват тази информация. Методологията на изследването се основава на метода „Същност – Връзка” (ER-модел или entity-relationship model).

В резултат на изследването е разработен концептуален модел, който представлява обобщен поглед върху „библиографския универсум”. Моделът е независим от отделни правила за каталогизация или от приложения на понятията, които представя самият той.

Експертите разглеждат библиографския запис като съвкупност от данни, свързани с обектите, описани в библиотечните каталози и националните библиографии. В тази съвкупност са включени описателни данни (определени в Международния стандарт за библиографско описание - ISBD), данни, използвани в точките за достъп (данни за лица, колективни органи, заглавия и предмети, които служат за инструменти при подреждане или за вход в показалци), други организиращи данни (като класификационни индекси, анотации, резюмета) и специфични данни за екземплярите в библиотечните колекции (като инвентарни номера и сигнатури).

Дефинираните за това изследване обекти, които са от основен интерес за потребителите на библиографските данни са представени в три групи:

⁴⁶ **Функционални изисквания към библиографските записи: Концептуален модел** [online]. [Прегледан 06.04.2017]. PDF. Достъпно от: https://www.ifla.org/files/assets/cataloguing/frbr/frbr_2009-bg.pdf; **Functional requirements for bibliographic records: Final report** [online]. München: K.G. Saur, 1998. As amended and corrected through February 2009. [Viewed 17.07.2017]. PDF. Available from: http://www.ifla.org/files/assets/cataloguing/frbr/frbr_2008.pdf

Обекти от първа група -

Продуктите на интелектуална или художествена дейност, същности, които се описват в каталозите:

произведение	work
проявление	expression
въплъщение	manifestation
единица	item

Обекти от втора група -

Отговорят за интелектуалното или художествено съдържание, за физическото изработване и разпространението или за съхраняването на тези продукти:

лице	person
колективен орган	corporate body

Моделът *Функционални изисквания към нормативните данни (FRAD)* разглежда и обекта:

фамилия	family
---------	--------

Обекти от трета група -

Служат като теми (предмети) на интелектуална или художествена дейност:

понятие	concept
предмет	object
събитие	event
място	place

Моделът *Functional Requirements for Subject Authority Data (FRSAD)* разглежда всички обекти от трите групи като тема/предмет на произведение:

тема	thema
------	-------

Моделът дава рамка, която да осигури ясно, точно формулирано и общоприето разбиране за това, каква информация трябва да предоставя библиографския запис и най-вече какво очакваме да се постигне чрез него в отговор на потребностите на ползвателя. Основните опции на ползвателите при търсене и използване на библиографска информация са:

- използване на данните за намиране на материали, които съответстват на заявените от ползвателя критерии за търсене (напр. при търсене на документи с определено съдържание или на звукозапис с определено заглавие);
- използване на извлечените данни за идентифициране на обект (напр. за да се потвърди, че описаният в записа документ отговаря на документа, търсен от ползвателя или за да се разграничат два текста или два звукозаписа с едно и също заглавие);
- използване на данните за подбиране на обект, който съответства на потребностите на ползвателя (напр. да се подбере текст на език, който ползвателят разбира или версия на компютърна програма, съвместима с хардуер и операционна система, достъпни за ползвателя);
- използване на данните за придобиване или получаване на достъп до описания обект (напр. поръчка за покупка на издание, искане за заемане на екземпляр от книга от библиотечната колекция или за онлайн достъп до електронен документ, съхраняван в отдалечен компютър).

В контекста на модела се отделя сериозно внимание на библиографските връзки между обектите. Обикновено ползвателят формулира запитване за търсене, като използва един или няколко атрибута на търсения от него обект, и чрез тях го открива. Отразените в библиографския запис връзки помагат на ползвателя да „навигира“ в съдържанието, представено в библиография, каталог или библиографска база данни и да установи корелации между намерения обект и други, свързани с него, обекти.

Моделът FRBR се превръща в изходна точка за многобройни последващи разработки на правила за каталогизация, комуникативни формати и различни системи, осигуряващи създаването, управлението и ползването на метаданни. Нещо повече, популярността на резултатите от изследването през последните години показва, че неговото въздействие значително надхвърля областта на библиографското описание, засягайки принципни въпроси на метаданните в дигитална среда. Възприемането на каталога като организирано множество от данни, описващи съдържанието на съхранявани информационни обекти, дава основание да се приложи същият концептуален модел не само върху формалното описание на материалите, но и върху точките за достъп към това съдържание при изследването на инструментите за контрол върху формите на имената, наименованията на колективни органи и предметните рубрики.

FRBRoo е „интерпретация“ на концептуалния модел FRBR, която разглежда обектите на интелектуална или художествена дейност като резултат на процесите създаване, реализация, планиране и др. FRBRoo версия 1.0 (публикувана през 2009 г.) е

разширена съгласно CIDOC CRM, така че да обхваща информация, свързана с музеите. През 2015 г. е одобрена версия 2.4⁴⁷, която включва и обектите, атрибутите и връзките, отразени във FRAD и FRSAD.

Подходът на FRBRoo позволява да се вземат под внимание обстоятелствата, при които конкретно произведение е било замислено или реализирано. В резултат обектите „произведение”, „проявление”, и „въплъщение” са разбити на класове със специфични свойства. Така във FRBRoo „произведение” (Work), напр., се разглежда като индивидуално произведение (Individual Work), съставно произведение (Complex Work), съпроводително произведение (Container Work), съвкупност от произведения (Aggregation Work), продължаващо произведение (Serial Work), публикувано произведение (Publication Work), изпълнителско произведение (Performance Work), записано произведение (Recording Work). Т. е. произведение се явява суперклас, обединяващ класове като частни случаи, всеки от които има своя специфика на създаване.

FRBRoo е онтология на библиографската информация като цяло. Тя обхваща всички библиографски ресурси, без обаче да взема предвид спецификите при описание на различни ресурси. Това е причината Международният център за ISSN (ISSN International Centre) да инициира специфицирането на моделите FRBRoo и CIDOC CRM за продължаващите ресурси. В резултат експерти от IFLA, Bibliothèque nationale de France, ISSN International Centre създават PRESSoo – концептуален модел на библиографска информация, свързана с продължаващи ресурси⁴⁸.

1.3.2.3. IFLA LRM

Концептуален модел IFLA Library Reference Model (IFLA LRM)⁴⁹. Освен по-общото изследване на библиографските данни, ИФЛА чрез своите специализирани работни групи „Функционални изисквания и номериране на контролните записи” (Functional Requirements and Numbering of Authority Records, FRANAR) и „Функционални изисквания към нормативните предметни записи” (Functional Requirements for Subject Authority Records,

⁴⁷ **Bekiari**, Chryssoula, **Doerr**, Martin, **Bœuf**, Patrick Le, **Riva**, Pat, eds. *Definition of FRBRoo: A conceptual model for bibliographic information in object-oriented formalism* [online]. Version 2.4. November 2015. [Viewed 24.05.2018]. PDF. Available from: https://www.ifla.org/files/assets/cataloguing/FRBRoo/frbroo_v_2.4.pdf

⁴⁸ **Bœuf**, Patrick Le, ed. *PRESSoo: Extension of CIDOC CRM and FRBRoo for the modelling of bibliographic information pertaining to continuing resources* [online]. Version 1.3. August 2016. [Viewed 24.05.2018]. PDF. Available from: https://www.ifla.org/files/assets/cataloguing/PRESSoo/pressoo_v1-3.pdf

⁴⁹ **Riva**, Pat, **Bœuf**, Patrick Le, **Žumer**, Maja, eds. *IFLA Library Reference Model: A Conceptual Model for Bibliographic Information* [online]. August 2017 Revised after world-wide review. [Viewed 19.08.2017]. PDF, 4.73 MB. Available from: <https://www.ifla.org/files/assets/cataloguing/frbr-lrm/ifla-lrm-august-2017.pdf>

FRSAR) предприема анализ на специфични аспекти на библиографския универсум като нормативните данни*. Анализират се обектите „име” и „заглавие” и предметните термини, в резултат на което са създадени: „Функционални изисквания към нормативните данни” (Functional Requirements for Authority Data, FRAD, 200950) и „Функционални изисквания към предметните нормативни данни” (Functional Requirements for Subject Authority Data, FRSAD, 201051). Задачата на трите т.нар. „FR-модели” [FRBR, FRAD, FRSAD] е да обхванат и анализират цялата библиографска система. Неизбежно обаче те се различават значително в обхвата и гледните си точки, както и в решенията, приети за някои общи въпроси. Ето защо експертите на ИФЛА предприемат следваща стъпка – създаване на консолидиран, съгласуван и пълен концептуален модел на библиографската информация (2010-2015). През февруари 2016 г. е публикуван проект на „FRBR-Library Reference Model” за преглед от световната професионална съобщност. Следващата година (март 2017 г.) излиза ревизирана публикация и с ново наименование на модела - IFLA Library Reference Model (IFLA LRM). На 17 август 2017 г. документът е публикуван след преглед в световен мащаб и предстои одобряване от Професионалния комитет или Управителния съвет на IFLA (Professional Committee or Governing Board)**.

IFLA LRM има за цел да бъде концептуален модел на високо ниво, разработен според метода за моделиране на данни „обект – връзка”. Този метод се основава на три основни понятия – „обект” (същност), „връзка” и „атрибут” (свойство). По дефиниция „обект” е всяка различима същност – човек, място, организация, събитие и др., за които е необходимо да се съхранява информация, „връзка” показва отношенията между същностите, а „атрибут” характеризира свойство на обект (възможно и връзка).

Моделът обхваща библиографските данни, разбирани в широк, по-общ смисъл. Методологията на процеса на моделиране довел до IFLA-LRM произтича от подхода, възприет в оригиналното изследване FRBR, където той е описан така: „Намерението е това изследване да очертае с ясно дефинирани термини функциите, които изпълнява

* Нормативните данни представляват контролирани точки за достъп и друга информация, която се използва за подреждане на произведенията на определено лице, фамилия или колективен орган или различните издания на едно заглавие. Контролираните точки за достъп включват установени и вариантни форми на името (наименованието), обединени за идентифициране на обект.

⁵⁰ **Функционални изисквания към нормативните данни: Концептуален модел** [online].

[Прегледан 24.05.2018]. PDF. Достъпно от: https://www.ifla.org/files/assets/cataloguing/frad/frad_2009-bg.pdf; **Functional requirements for authority data: A conceptual model** [online]. As amended and corrected through July 2013. [Viewed 24.05.2018]. PDF. Available from: http://www.ifla.org/files/assets/cataloguing/frad/frad_2013.pdf

⁵¹ **Functional requirements for subject authority data (FRSAD): A conceptual model** [online]. [Viewed 17.07.2017]. PDF. Available from: <http://www.ifla.org/files/assets/classification-and-indexing/functional-requirements-for-subject-authority-data/frsad-final-report.pdf>

** Информация към 24.05.2018 г. на <https://www.ifla.org/publications/node/11412>, Last update: 14 March 2018

библиографския запис, като взема предвид различни носители, различни приложения и различни потребности на ползвателите. Изследването трябва да обхване пълния диапазон от функции на библиографския запис в най-широк смисъл, т.е. запис, който съдържа не само описателни елементи, но и точки за достъп (име, заглавие, предмет и др.), други „организационни“ елементи (класификация и др.) и анотации.”

IFLA LRM преразглежда и дефинира всички потребителски задачи, обекти, атрибути и връзки, свързани с библиографската информация. Семейната модели FRBR, както и LRM, започват от потребителските задачи, които трябва да бъдат активирани и поддържани от библиографските информационни системи. Тези потребителски задачи осигуряват границите на модела и служат като отправна точка за дефиниции на обекти, атрибути и връзки. Библиографските и нормативните данни представляват интерес за широка група потребители – от библиотечни потребители (читатели, изследователи, студенти и др.) до библиотекари и други членове на информационната верига, като издатели и книгоразпространители. Тези потребителски групи имат различни нужди и различни приоритети. LRM следва FRBR при избора на основната потребителска група: крайни потребители и библиотекари. Библиотекарите, които създават и поддържат метаданни, понякога могат да изпълняват същите задачи като част от своята работа.

Потребителски задачи (User Tasks)

IFLA LRM дефинира пет обобщени потребителски задачи, свързани с библиотеките: намиране (Find), идентифициране (Identify), подбиране (Select), получаване (Obtain), изследване (Explore) (Табл. 2).

Таблица 2. IFLA LRM Потребителски задачи.

Потребителски задачи	
Намиране	Откриване/Предоставяне на информация за ресурс/и, представляващи интерес за потребителя чрез търсене по подходящи критерии.
Идентифициране	Разграничаване на отделни ресурси сред откритите ресурси.
Подбиране	Определяне/Преценяване пригодността на намерените ресурси и възможността да се приемат или отхвърлят конкретни ресурси.
Получаване	Осигуряване на достъп до ресурс/и.
Изследване	Откриване на ресурсите, използвайки връзките между тях.

Първите четири задачи се дефинират като обобщения на задачи във FRBR, FRAD и FRASAD със същото наименование. „Изследване” е взимствана от FRASAD, но тук е определено, така че да включва аспекти, извлечени от контекста на задачата

„контекстуализиране” във FRAD. Тя се свързва с възможността потребителят да проследява връзката между два или повече обекта (напр. връзката между лице и име). За реализирането на тази задача е необходимо изграждането на ясни връзки между обектите, което би създавало условия за получаване на контекстуална информация и възможност за навигация.

Обекти (Entities)

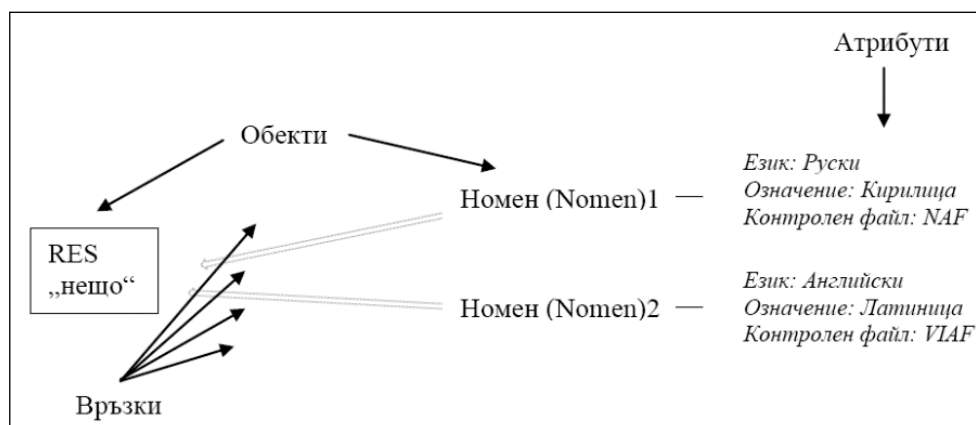
За разлика от FR-моделите, където се наблюдава „плоска” структура, в IFLA LRM обектите, атрибутите и връзките са представени в йерархия.

IFLA LRM определя 11 обекта: „нещо” (res – лат. ез. = „thing” на англ. ез.), „произведение” (work), „проявление” (expression), „въплъщение” (manifestation), „единица” (item), „агент” (agent), „лице” (person), „колектив” (corporate body), „име” (nomen), „място” (place), „период от време” (time-span). Моделът включва единствен обект от най-високо ниво – „нещо” (res). Всички други обекти са директни или индиректни негови подкласове. Осем от обектите са йерархично обвързани с „нещо” – „произведение” (work), „проявление” (expression), „въплъщение” (manifestation), „единица” (item), „агент” (agent), „име” (nomen), „място” (place), „период от време” (time-span). Останалите два обекта са подчинени на обекта от второ ниво на йерархията - „агент”. Йерархията на обектите и техните определения в референтния модел IFLA LRM са илюстрирани с примери и са представени в Приложение 1.

Атрибути (Attributes)

Докато обектите са опорните точки, рамката на модела, а връзките свързват обектите един с друг, то атрибутите описват (характеризират) обектите (фиг. 6).

Фигура 6. Елементи на IFLA LRM



Съществуват два начина за изразяване на атрибут⁵²:

– чрез буквен или числов низ, литерал за атрибути на данни (Datatype properties), съгласно разграничението на OWL (Web Ontology Language), дефинират се чрез RDF literals и XML Schema datatypes;

– чрез URI (Uniform Resource Identifier), посочващ външен източник (авторитетен или нормативен документ, например файл с нормативни записи или списък с кодирани стойности) за атрибути на обекти (Object properties) според терминологията на OWL.

IFLA LRM включва 24 атрибута, представени в йерархичен порядък. Един и същ атрибут може да принадлежи на различни същности, напр. атрибутът „Права за използване“ описва проявление, възплъщение, единица. Атрибутите се отнасят до десет от обектите, изключение прави обектът „колектив“, за който не са посочени атрибути. Примери за „атрибути“ включват: Условия за достъп, Картографски мащаб, Език, Вид на носителя, Информация за контакти, Права за използване, Име, Читателско предназначение, Схема, Източник и др. Списъкът с атрибутите, изброени в LRM, не е изчерпателен, ако е необходимо могат да се добавят допълнителни атрибути, взети от RDA (Resource Description and Access)⁵³.

Връзки (Relationships)

IFLA LRM описва 36 връзки, номерирани последователно от LRM-R1 до LRM-R36. Връзките в модела са реципрочни, например:

(LRM-R2) Произведение „се реализира/осъществява чрез“ Проявление <WORK ‘is realized through’ EXPRESSION>

(LRM-R2i) Проявление „реализира“ Произведение <EXPRESSION ‘realizes’ WORK>

Изразени в графичен вид, обектите и връзките в IFLA LRM са представени на фиг. 7. За да се избегне претрупаност, не се посочва йерархичната връзка „явява се“ („IsA“), която свързва всички обекти с горестоящия обект „нещо“ (res), а се означават само еднопосочните връзки, илюстриращи отношения между същностите.

⁵² *IFLA Library Reference Model: A Conceptual Model for Bibliographic Information* [online], August 2017 Revised after world-wide review, p. 11. [Viewed 19.08.2017]. PDF, 4.73 MB, Available from: <https://www.ifla.org/files/assets/cataloguing/frbr-lrm/ifla-lrm-august-2017.pdf>

⁵³ *RDA toolkit* [online]. Available from: <http://www.rdatoolkit.org/>

1.3.2.4. CERIF

Общ европейски формат за изследователска информация CERIF (Common European Research Information Format)⁵⁵. CERIF е модел, който концептуализира научноизследователската сфера. Той има 2 нива на приложение: логическо ниво – като модел, който описва същностите (обектите) и връзките между тях в областта на научните изследвания; физическо ниво – като концептуална схема на база от данни (Database Scripts)⁵⁶, т.е. служи за проектиране на конкретни бази от данни. CERIF обхваща всички аспекти на научноизследователската информация – проекти, лица, организации, патенти, продукти, библиометрия, идентификатори и др.

Общият европейски формат за изследователска информация CERIF се прилага като:

а) международна техническа препоръка за съхраняване и съвместимост на научноизследователска информация на Европейския съюз към страните-членки. Той се поддържа от организация *euroCRIS* (www.eurocris.org), чиято дейност е посветена на разработването на информационни системи за научни изследвания (Current Research Information Systems, CRIS);

б) референтен модел за създаване на конкретни системи за научни изследвания (Current Research Information Systems, CRIS);

в) стандарт за обмен на данни (CERIF-XML) между системи;

г) средство за осигуряване на единен достъп до хетерогенни системи.

CERIF е резултат от проект на Европейската комисия, осъществен в два етапа - 1987–1990 г. и 1997–1999 г. От неговото създаване през 2004 г. до днес моделът CERIF се развива постоянно. През 2006 г. излиза сериозно актуализирана версия (CERIF-1.1 FDM) с добавен така нар. „семантичен слой” (класификация на обектите) и с включена спецификация за обмен на данни във XML формат. Последната модификация на модела е CERIF 1.6 FDM от юли 2013 г., която включва и модул CERIF 1.6 XML Schema файл⁵⁷.

Създаването на общоприет концептуален модел за научноизследователска информация е важен резултат в технологиите за семантично структуриране на съдържание, създавано в академичните и научни среди, както и в научните електронни библиотеки в частност. Моделът CERIF поддържа управлението на научноизследователска информация

⁵⁵ CERIF. In: *euroCRIS Current Research Information Systems*. [online]. [Viewed 24.05.2018]. Available from:<https://www.eurocris.org/cerif/main-features-cerif>

⁵⁶ Ibid.

⁵⁷ CERIF-1.6. In: *euroCRIS Current Research Information Systems* [online]. [Viewed 24.05.2018]. Available from:<http://www.eurocris.org/cerif/feature-tour/cerif-16>

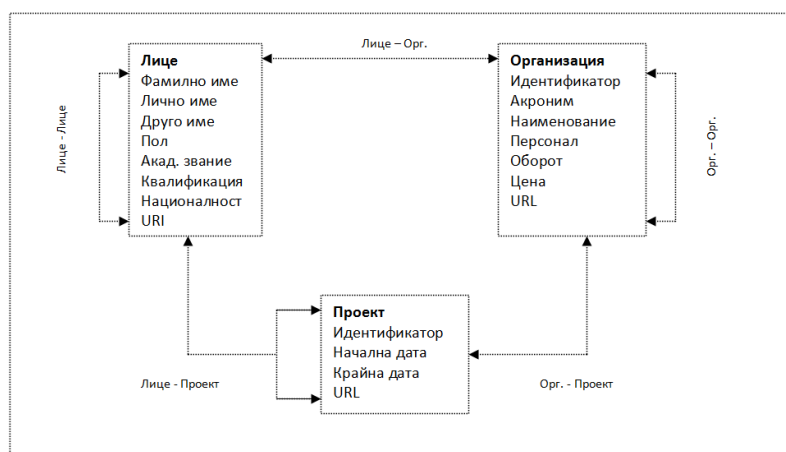
и прави възможно взаимодействието между информационни системи, свързани с науката. CERIF дава схема, която гарантира оперативна съвместимост.

Моделът има следните основни характеристики:

- представлява модел на предметна област, по-конкретно формално описва аспекти на изследователската сфера;
- построен и въз основа на метода „Същност-Връзка”, отразява атрибутите на обектите и на връзките между тях, което дава богата семантика;
- поддържа многоезичност;
- използвани са неограничен брой контролирани речници (тезауруси, таксономии, класификации и др.);
- разширяем е, което предполага съвместимост и го прави приложим за обмен на данни в хетерогенни разпределени среди. Използването на т. нар. семантичен слой осигурява гъвкавост и мащабируемост.

Моделът е доста сложен. С последно добавените елементи (модификация 2013) съдържа 293 обекти, 1814 атрибути и 665 връзки. Обектите са представени в 5 групи. С най-голяма тежест са обектите от първата група, която включва: Лице (Person), Организация (OrganisationUnit) и Проект (Project)*. Тези базови обекти на модела CERIF представляват участниците (лица и научни организации) и научно-изследователската дейност (проект). Основните обекти влизат в рекурсивни и свързващи връзки (фиг.8).

Фигура 8. Обекти и ивръзки в CERIF



Като важно предимство на CERIF се сочи неутралността към архитектурата - може да бъде реализиран чрез различни модели на данни (релационен, обектно-ориентиран,

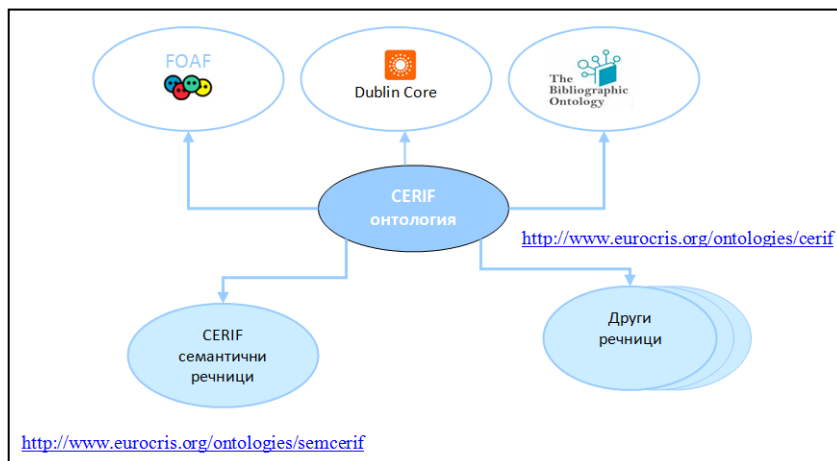
*Само в модификацията на модела от 2006 г. като основен обект се включва „ResultPublication”.

текстово търсене, уеб) и да бъде реализиран като електронна библиотека, система за управление на бази от данни, уеб страница, технологии за управление на знания и др.

През 2012 г. CERIF е развит в онтологична структура⁵⁸. Тя адаптира модела CERIF в съответствие с парадигмата „свързани данни” (Linked Data), така че семантичните уеб приложения могат да използват тези термини в различни формати и приложения, съвместими с RDF. Следващата година излиза актуализирана версия (Version 0.2 February 2013). В нея атрибутите са преобразувани в онтологични аксиоми, а обектите са преобразувани в RDF класове и техните свойства (26 класа и 58 свойства*).

Онтологията използва контролиран, машинночетим речник в RDF - CERIF Semantic Vocabulary за съответните типове обекти и роли в изследователския контекст. Речникът включва RDF класове, които описват напр. типове организации или ролята на лице в изследователски проект. Онтологията използва и „външни” речници, когато е необходимо да опише някои класове обекти: The Bibliographic Ontology (<http://bibliontology.com>), DCMI Metadata Terms (<http://dublincore.org/documents/dcmi-terms/>), Friend of a Friend Vocabulary (FOAF Vocabulary Specification, <http://xmlns.com/foaf/spec/>)⁵⁹ (фиг. 9)

Фигура 9. CERIF взаимодействие с онтологии



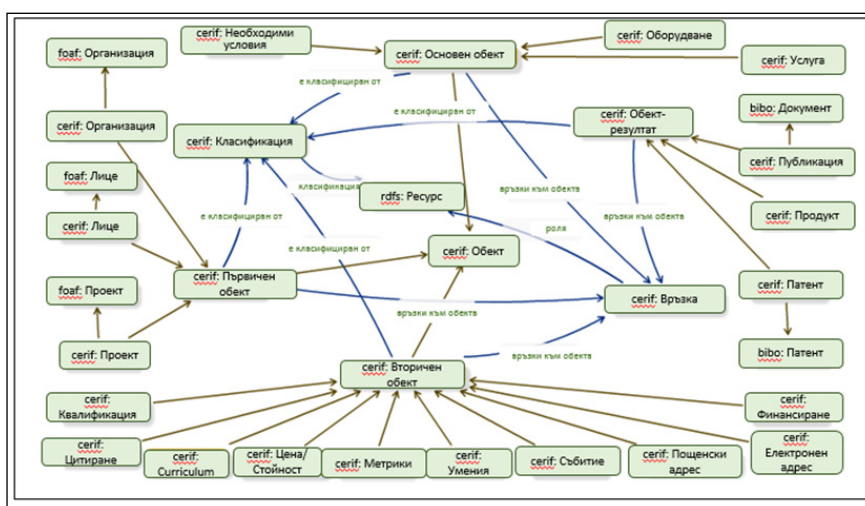
⁵⁸ Sicilia, Miguel-Ángel. *The Common European Research Information Format Ontology CERIF Ontology 0.2* [online]. Namespace Document 01 February 2013. Vocabulary based on CERIF Model 1.3 [Viewed 24.05.2018]. Available from: <http://eurocris.org/ontologies/cerif/1.3/index.html>

* Вж.: <https://www.eurocris.org/ontology>

⁵⁹ *The Common European Research Information Format CERIF Semantic Vocabulary 0.2*. [online]. Namespace Document 01 February [2013]. Vocabulary based on CERIF Model 1.5. [Viewed 24.05.2018]. Available from: <https://www.eurocris.org/ontologies/semcerif/#sec-glance>

Обобщен изглед на онтологията на информация за научните изследвания CERIF е представена на фиг.10. Изобразени са 26-те класове обекти в CERIF и взаимстваните от: RDFS - rdfs:Resource, от онтологията FOAF - foaf:Organisation, foaf:person, foaf:project и от BIBO bibo:document.

Фигура 10. Онтология на информация за научните изследвания CERIF⁶⁰.



Организацията euroCRIS, която поддържа CERIF, е ангажирана със създаването на условия за съвместимост между системите за научна информация. Във връзка със своята дейност по проект „Инфраструктура за отворен достъп до научни изследвания в Европа” (Open Access Infrastructure for Research in Europe, OpenAIRE, финансиран от Седма рамкова програма за научни изследвания и технологично развитие и Европейския съвет за научни изследвания (European Research Council, ERC) тя публикува през 2015 г. документа „OpenAIRE Guidelines for CRIS Managers”^{*}. „Насоките” имат за цел да направят възможна съвместната работа на съществуващите информационни системи за научни изследвания (CRIS) и инфраструктурата на OpenAIRE. Обменът на информация се основава на модела за данни на общия европейски изследователски информационен формат (CERIF) в XML обменен формат и протокола OAI-PMH. Насоките са предназначени за всички администратори на системи, които предоставят информация за научните изследвания на отворен достъп.

⁶⁰ Източник: The **Common** European Research Information Format CERIF Semantic Vocabulary 0.2 ...

* Достъпно от: <https://www.eurocris.org/openaire-gl>

1.4. Оперативна съвместимост на електронните библиотеки

Едно от основните направления на изследванията на електронните библиотеки е осигуряването на оперативна съвместимост. Терминът „оперативна съвместимост“ (interoperability) описва способността и инструментите за пряка комуникация и сътрудничество между различни системи, основани на общи стандарти, технологии и концепции. Оперативната съвместимост в трите ѝ аспекта: структурен, семантичен и синтактичен се свързва със способността на една система да комуникира и работи с друга (обикновено от различен тип) с цел ефективен обмен и използване на информация⁶¹.

До момента налице са голям брой теоретични и приложни разработки за оперативна съвместимост, но въпреки това липсва единствен възприет от всички подход. Комуникацията може да има различна форми, напр. прехвърляне, обменяне, трансформиране или интегриране на информация. Когато системите могат да комуникират взаимно, то е налице *синтактична оперативна съвместимост* (syntactic interoperability). Синтактичната оперативна съвместимост описва точния формат на информацията, която се обменя по отношение на граматика, формат и схеми. Възможността за достатъчно точно автоматично интерпретиране на значението на обменяната информация в съответствие с очакванията на крайните потребители е критерий за *семантична оперативна съвместимост* (semantic interoperability)⁶². Семантична оперативна съвместимост между две или повече електронни библиотеки е способността им да „разбират“ значението на обменената информация по един и същи начин. Това, което прави семантичната оперативна съвместимост сложна, е фактът, че начинът, по който се използват софтуерните системи, техните цели и контекст, е различен. Така предоставянето на данни, кодирането и вариациите на значенията също са различни. В раздел D5.3.1 „Семантична оперативна съвместимост за системите на електронна библиотека“ (Semantic Interoperability in Digital Library Systems) на проектът DELOS се разглеждат теоретичните предпоставки за осигуряване на семантична оперативна съвместимост. Като такива се посочват стандарти, консенсуси, базови онтологии, системи за организация на знанието, семантични услуги,

⁶¹ **Interoperability**. In: Reitz, Joan M. *ODLIS – Online Dictionary for Library and Information Science* [online]. [Viewed 24.05.2018]. Available from: http://www.abc-elio.com/ODLIS/odlis_i.aspx

⁶² Нишева-Павлова, Мария. *Създаване на хетерогенни цифрови библиотеки...* [онлайн]. с. 24.

архитектура и инфраструктура, както и въпроси, свързани с достъпа и правата⁶³. И все пак въпросът, как те трябва да бъдат приложени на практика, остава без отговор⁶⁴.

Проблемите със семантичната оперативна съвместимост са свързани с всички аспекти от жизнения цикъл на документите и информацията: създаване и модифициране (промяна); публикуване; комплектуване, подбор, съхраняване, изграждане на колекции; каталогизиране (метаданни за описание, иидентификация, получаване на идентификатор, регистриране); индексирание; интегриране и организиране; достъп, търсене и откриване; архивиране и съхранение и др.⁶⁵

Постигането на семантична оперативна съвместимост между системите на различни електронни библиотеки е от съществено значение за следните основни аспекти:

- 1) осигуряване на възможности за формулиране и изпълнение на сложни информационни запитвания;
- 2) подобряване точността на търсене;
- 3) интегриране на хетерогенни информационни ресурси.

Оперативната съвместимост е възможна чрез прилагането на стандарти⁶⁶. Постигането на семантична оперативна съвместимост между електронни библиотеки пряко зависи от процесите на стандартизация. Чрез тях се създават условия добрите практики и постижения да се превърнат в норма. Стандартизация в областта на семантичната оперативна съвместимост на електронните библиотеки се налага в следните главни направления:

- 1) унифициране и стандартизиране на формата и съдържанието на схемите за описание на цифровото съдържание,
- 2) стандартизиране в идентификацията на понятия и обекти от реалния свят (създаването на идентификатори).

⁶³ Patel, M., Koch, T., Doerr, M., and Tsinaraki, C. *DELOS: A Network of Excellence in Digital Libraries* [online]. D5.3.1: Semantic Interoperability in Digital Library Systems. Task 3: Semantic Interoperability WP5: Knowledge Extraction and Semantic Interoperability, July 2004–June 2005. [Viewed 30.05.2018]. PDF. Available from: <http://delos-wp5.ukoln.ac.uk/project-outcomes/SI-in-DLs/SI-in-DLs.pdf>.

⁶⁴ Olensky, Marlies. Semantic interoperability in Europeana. An examination of CIDOC CRM in digital cultural heritage documentation. In: *Bulletin of IEEE Technical Committee on Digital Libraries* [online]. 2010, Vol. 6, Issue 2 [Viewed 30.05.2018]. Available from: <http://www.ieee-tcdl.org/Bulletin/v6n2/Olensky/olensky.html#18>

⁶⁵ Patel, Manjula, Koch, Traugott, Doerr, Martin, Tsinaraki, Chrisa. Semantic Interoperability in Digital Library Systems [presentation]. In: *DELOS WP5 Workshop: Semantic Interoperability in DL systems* [online]. 17th September 2004, Bath, UK. [Viewed 24.05.2018]. Available from: http://opus.bath.ac.uk/35098/4/WP5_Workshop_040916.pdf.

⁶⁶ **Interoperability**. In: *Standards Glossary* [online]. [Viewed 10.06.2017]. Available from: https://www.ieee.org/education_careers/education/standards/standards_glossary.html.

Използването на единни стандарти прави възможно предаването и интегрирането на информация без междинни трансформации. Системният подход към семантичната оперативна съвместимост на електронни библиотеки предполага тристепенна схема на осигуряване: стандарти за метаданни, общи онтологии, предметни онтологии⁶⁷.

Съществува специален международен стандарт, приет и в България, даващ рамка, която дефинира елементите за описание на документите. Прилагането на стандарта има за цел да:

- а) позволи стандартизирано описание на документи и обекти,*
- б) установи общоприето разбиране за фиксирани нива на съвкупност, за да се осигури оперативна съвместимост на документите и информацията за тях, между организационните системи,*
- в) позволи вторична употреба и стандартизиране на метаданните за управление на документи във времето, пространството и чрез различни софтуерни приложения⁶⁸.*

Стандартизацията на метаданните има връзка с използването на онтологии. Онтологиите служат като речници на термините, с които се формулират метаданните, т.е. източници на лексиката на метаданните⁶⁹. Обикновено създаването на една предметна онтология има за цел да се дефинира значението на понятията и термините, използвани в рамките на определена предметна област. Така използването и споделянето на подходящи онтологии помага да бъде постигнато общо съгласувано между софтуерните системи на електронни библиотеки разбиране за съдържанието на съхраняваните информационни ресурси⁷⁰. Същността и използването на онтологии в електронните библиотеки се разглежда в Глава 2.

1.5. Инструментални средства за създаване на електронни библиотеки

За създаване на електронни библиотеки и за осигуряване на достъп (възможно и други услуги) до тяхното съдържание се използват различни програмни пакети.

За да установим кои са най-използваните от тях проведохме проучване в Директорията на хранилищата с отворен достъп OpenDOAR (The Directory of Open Access Repositories*). В таблица 3 са представени 15 от най използваните в света програмни

⁶⁷ Нишева-Павлова, Мария. *Създаване на хетерогенни цифрови библиотеки...* [онлайн]. с. 25.

⁶⁸ *БДС ISO 23081-2:2015 Информация и документация. Управление на метаданни за документи/записи. Част 2: Концептуални въпроси и прилагане.* София: БИС, с. 4.

⁶⁹ Нишева-Павлова, Мария. *Създаване на хетерогенни цифрови библиотеки ...* [онлайн]. с. 25.

⁷⁰ Пак там, с. 26.

* <http://www.opendoar.org/index.html>

продукти. Според критериите програмите да бъдат с отворен код и със световно разпространение бяха подбрани 3 продукта. Освен това ще представим и допълнителен такъв, използван в България.

Таблица 3. Използвани програмни инструменти за електронни библиотеки (ЕБ) с отворен достъп (ОАІ-РМН) и разпространението им**

Название на софтуер	Брой ЕБ (ОАІ-РМН)	Разпространение	Пример за ЕБ и брой записи	
Dspace	1525	световно	DML-CZ Czech digital mathematics library, Чехия	38337
Eprints	469	световно	Digital Library NAPS of Ukraine, Украйна	13312
Digital Commons	164	САЩ, Канада, Австралия	DAFWA Research Library, Австралия	4478
OPUS	85	Германия, Сърбия	Universitätsbibliothek Bamberg, Германия	42784
dlibra	60	Полша	Jagiellonian Digital Library, Полша	378243
CONTENTdm	56	САЩ	USC Digital Library University of Southern California, United States	437978
HAL	56	Франция	Centre pour la communication scientifique directe, France	1418291
Greenstone	50	световно	Papers Past National Library of New Zealand	50474317
Fedora	49	САЩ, Европа, Австралия	NSDL (National Science Digital Library), САЩ	271280
HTML	35	световно	DSAL (Digital South Asia Library), САЩ	
Open Repository	31	световно	University of Arizona (UA), САЩ	62284
DiVA-Portal	30	Швеция	Enheten för digital publicering, Uppsala Universitet, Швеция	324756
Islandora	24	Хърватия, Канада, САЩ	University of Prince Edward Island (UPEI), Канада	250520
Digibib	23	Испания	Biblioteca Digital AECID, Испания	8248
Invenio	21	Европа	Organisation Européenne pour la Recherche Nucléaire (CERN), Швейцария	322352

Установено бе, че най-използваните програмни продукти със свободно използване и с широко световно разпространение са DSpace, EPrints, Greenstone. Dspace и Eprint са

** Справката е от 12.01.2018 г.

предназначени за създаване преди всичко на институционални репозиториуми, а Greenstone е програмен продукт за електронни библиотеки.

DSpace

Първата публична версия на DSpace е пусната през ноември 2002 г. като резултат от съвместната дейност между изследователи от библиотеката на Масачузетския технологичен институт (Massachusetts Institute of Technology, MIT) и HP Labs. Днес той се поддържа от организацията с нестопанска цел DuraSpace и се разпространява и използва свободно под лиценз BSD (Berkeley Software Distribution license).

DSpace представлява платформа за дългосрочно съхранение на цифрови документи на академични и научни организации и функционира като централизирано информационно хранилище на организацията. Отделните подразделения в структурата на организацията (факултети, катедри, лаборатории, центрове, школи, отдели или др.) са самостоятелно представени в рамките на системата. Всеки потребител еднолично или чрез оторизирано лице, което да преглежда и редактира метаданните преди включването им в основния репозиториум, има възможност непосредствено да депозира документи чрез потребителския уеб-интерфейс.

Някои от функциите на DSpace, напр. навигиране и търсене на документи, могат да се изпълняват анонимно (без регистрация), за други - като добавяне на документ, е необходима потребителска регистрация.

За всеки регистриран потребител DSpace съхранява лична информация като адрес на електронна поща, име и фамилия, парола; отразява индивидуалните му предпочитания и интереси, като поддържа списък с колекции, за които потребителят желае да получава уведомления при нови постъпления.

DSpace разполага с развита система от потребителски права (роли) в сравнение с други програмни пакети, напр. EPrints. Според модела на данните потребителите могат да имат различни роли: автори, администратори, потребители на съдържание, оторизирани потребители, които имат достъп до отчетна информация и др.

Търсене и навигиране (преглеждане, Browse)

DSpace позволява преглеждане на документите в електронната библиотека по заглавие, автор, дата на публикуване, отдели (според структурата на организацията, напр. факултети, катедри и др.), колекции. Платформата предоставя възможност за основно търсене (по дума от описанието или от пълния текст) и разширено търсене с определяне на признаците (елементи от библиографското описание, описателните метаданни) за търсене

и тяхното комбиниране чрез логически оператори. Възможно е търсене по автор, заглавие, идентификатор, език, предмет, спонсор (издател), резюме и др. Признаците, по които се провежда търсенето, могат да бъдат конфигурирани от администратора според интересите на потребителите.

В българските академични институции има натрупан опит в използването на DSpace – например Репозиториум на Института по математика и информатика на БАН*, Научен портал на Софийски университет „Св. Климент Охридски“**, Научен портал на Бургаския свободен университет***, Електронен репозиториум на Централна медицинска библиотека на Медицински университет - София****.

EPrints

EPrints е безплатен софтуер с отворен код за изграждане на хранилища с отворен достъп, които са в съответствие с инициативата Open Archives Initiative Protocol за събиране на метаданни. Той е близък по функционалност до системите за управление на съдържание⁷¹, но се използва главно за институционални хранилища и научни списания.

EPrints е разработен в резултат на проект TARDIS (2002–2008), част от програмата „Фокус върху достъпа до институционални ресурси“ (Focus on Access to Institutional Resources Programme) на Великобритания. Крайната цел на проекта е постигната – „създаване на устойчиво мултидисциплинарно институционално хранилище на Университета в Саутхемптън“ (University of Southampton) (<https://eprints.soton.ac.uk/>)⁷². Програмният продукт се разпространява под лиценз „Общо право на обществено ползване“ (GPL или GNU General Public License)⁷³.

EPrints специфицира три категории потребители: редактори, администратори и крайни потребители. В зависимост от „ролята“ си потребителят има достъп до повече или по-малко функции. EPrints предлага различен интерфейс на страницата според типа потребител. Потребителят се идентифицира с име и парола.

Чрез веб интерфейса на администратора могат да се изпълняват професионални дейности, свързани с управлението на ресурсите (електронни документи, метаданни,

* <http://sci-gems.math.bas.bg:8080/jspui/>

** <https://research.uni-sofia.bg/>

*** <http://research.bfu.bg:8080/jspui/>

**** <http://nt-cmb.medun.acad.bg:8080/oai/request>

⁷¹ Камбуров, Кирил. *Сравнителен анализ на онлайн платформи за електронни периодични издания: Дисертация*. София, 2018, с. 78.

⁷² TARDIS *University of Southampton e-Prints* [online]. [Viewed 24.05.2018]. Available from: <http://tardis.eprints.org/>

⁷³ EPrints [online]. In: *Wikipedia*. [Viewed 24.05.2018]. Available from: <https://en.wikipedia.org/wiki/EPrints>

потребители и др.), напр.: въвеждане и редактиране на записи, редактиране и актуализиране на предметния класификатор и др.

Потребителите, работещи с права на редактор, могат да преглеждат записите във временния буфер, да ги редактират или връщат за доработка. Само след потвърждаването (верифицирането) на записите те стават видими за крайните потребители на библиотеката.

Крайният потребител може да депозира документ и да въвежда запис само на своята „страница”. При създаването на нов запис за депозирания документ потребителят трябва да укаже типа/вида на документа. В EPrints се разграничават следните типове документи: статия от списание или вестник; книга или том от продължаващ сборник; съставна част от книга (раздел, глава); монография, технически отчет, отчет по проект, техническа документация (ръководство, протокол); доклад от конференция или семинар (или постер, слово, лекция, презентация); набор от данни; обучителни ресурси (лекция, упражнения, програми и др.) и други (вид документ, който не може да се включи в останалите категории). Типът документ предполага генериране на съответни елементи на описание. Само след попълването на необходимите библиографски данни, записът се прехвърля във временния буфер, където остава, докато не бъде обработен и потвърден от редактор.

Търсене и навигиране

Търсещата машина на EPrints индексира всички файлове (включително метаданни) в архива, което позволява реализирането на базово търсене и разширено търсене. Базовото търсене (simple search) представлява лексикално търсене в едно поле по дума от пълния текст или от библиографското описание. Разширеното търсене (advanced search) предоставя възможност за търсене по заглавие, автор, редактор, думи от резюмето, дата, ключови думи, предметни рубрики, тип на документа и други критерии, както и чрез логически оператори. Търсенето може да се ограничи по тип на документа, статус на документа (публикуван, непубликуван, под печат и др.), тип на информацията (текст, аудио и пр.) (Приложение 3)

Резултатите от търсенето могат да се сортират (ранжират) по име на автора, заглавие или година на издаване.

EPrints позволява навигиране, преглеждане на документите в електронната библиотека по дата на публикуване, автори, предметни рубрики, структурните подразделения на организацията (факултети, катедри и др. под.).

След създаването на електронна библиотека с програмен пакет EPrints е възможно задаването на допълнителни настройки, свързани с езика на интерфейса, с използването на метаданните и др.

В България два от осемте регистрираните научни институционални архиви в OpenDOAR са базирани на EPrints – Научен електронен архив на Нов български университет* (New Bulgarian University Scholar Electronic Repository)⁷⁴ и Академичен архив на Медицински университет – Варна (Medical Academic Repository)**.

EPrints и Dspace са програмни системи от един и същи клас. Те дават пълен спектър от функционалности за създаване и поддържане на архив (репозиториум) на научноизследователската продукция на научна или академична организация. Системите поддържат протокол OAI-PMH (The Open Archives Initiative Protocol for Metadata Harvesting) за събиране на метаданните на записите в архива, така че услугите могат да бъдат изградени с помощта на метаданни от много архиви. Осъществяването на OAI-PMH се реализира чрез използването на стандарт Dublin Core (ISO 15836:2009***). Двата програмни пакета се различават по структурно-логическото представяне на данните.

Организационният модел на данните в DSpace следва структурата на организацията, напр. факултети, ако става въпрос за университет. Разделът е най-високото ниво на йерархия в DSpace. Разделите могат да съдържат подраздели. Разделите или подразделите съдържат колекции от логически свързани документи.

Дадена колекцията принадлежи само на един раздел. Колекцията се състои от отделни документи, които представляват основната единица на архивиране. Всеки документ принадлежи само на една колекция, но допълнително може да бъде отразен и в друга колекция или колекции. Всеки документ (може да е 1 файл или файлове) има библиографско описание (елементи от метаданни), които служат за идентификация, навигация и търсене. Документите се описват по стандарт за описание Dublin Core (Qualified Dublin Core, QDC). Този запис по QDC е задължителен за всеки документ (не отделен файл). Той гарантира възможност за съвместимост, интегриране и намирането на документа. Метаданните QDC могат да бъдат въвеждани от потребители или да бъдат получени или извлечени от данни (ingest process).

В програмния пакет EPrints организационният модел е различен. Липсва строго структурно-логическо деление по раздели (отдели и подотдели) на организацията и по

* <http://eprints.nbu.bg/cgi/oai2>

⁷⁴ **Тодорова**, Радостина. *Свободен достъп до научни електронни публикации в Нов български университет* [online]. [Прегледан 23.05.2018]. Достъпно от: https://www.researchgate.net/publication/271450395_Svoboden_dostp_do_naucni_elektronni_publicacii_v_Nov_blgarski_universitet_6_Open_access_to_scientific_e-publications_in_the_New_Bulgarian_University

** <http://eprints.mu-varna.bg/cgi/oai2>

*** **БДС ISO 15836-1:2017** *Информация и документация. Множество от елементи на метаданни DUBLIN CORE. Част 1: Основни елементи. София: БИС, 2017*

колекции, както при DSpace. Организационната йерархия на институцията при представянето на научноизследователската продукция е важна при всеки институционален архив (репозиториум), затова софтуерът предполага възможност за навигация по отдели и подотдели на организацията. Важна характеристика на EPrints е преглеждането на съдържанието в библиотеката по предметна класификация. В продуктът е вградена широко използваната предметна класификация на Конгресна библиотека, но може да се създаде и специална класификация, която да отразява тематиката на конкретния електронен фонд на библиотеката.

Greenstone

Greenstone е многоезичен софтуерен инструмент с отворен код (GNU General License), предназначен за създаване на електронни библиотеки в интернет или на физически носители CD-ROM, DVD, флаш памет. Софтуерът е създаден през 1997 г. по проект за изграждане на електронна библиотека на Нова Зеландия, в който участват Университет в Уайкато (University of Waikato, New Zealand), Организацията на Обединените нации за образование, наука и култура – ЮНЕСКО и неправителствената белгийска организация „Human Info NGO”. Проектът има за цел реализацията на свободен софтуер, с който образователни и научни институции, институции на паметта и други обществени организации да изградят свои собствени електронни библиотеки⁷⁵. Електронна библиотека, създадена чрез програмния продукт Greenstone, може да съдържа множество колекции. Всяка колекция се съставя като съвкупност от документи в различни формати (Microsoft Word, Excel, Rich Text Format, HTML, plain text, PDF, ZIP, MP3), свързани въз основа на определен критерий. Към колекцията се прилагат единни механизми за съхраняване, индексирание, търсене, преглеждане и представяне.

Колекцията може да съдържа стотици хиляди, дори милион документи с различна по тип информация: текст (статии, книги, списания и др.п.), аудио и видео-материали. Колекциите могат да се структурират в подколекции или по логически обединения. Всяка колекция има собствен запис с параметрите, свързани със създаването и използването ѝ.

Текстовите документи могат да бъдат представени съдържателно по раздели (sections) – раздели, подраздели и подподраздели и т.н., които отразяват съдържателната структура на документа. Всеки от разделите от своя страна се състои от няколко абзаца (paragraphs).

⁷⁵ *Greenstone digital library software* [online]. [Viewed 24.05.2018]. Available from: <http://www.greenstone.org/>

Структурата на документа може да се използва при търсене. В резултатите се посочва в коя част от структурата се срещата търсената дума, напр. предговор. Ако документът няма вътрешна структура, то той се представя и преглежда във вид на последователни страници.

Информационните ресурси, които изграждат колекциите могат да се разполагат на локален компютър, в локална мрежа, а в глобалната мрежа да са достъпни чрез протоколи HTTP и FTP. Попълването на документи в библиотеката става не само локално от компютъра, но и чрез „сваляне” на файлове от мрежа. Софтуерът поддържа протоколи за обмен на данни между библиотеки, напр. Z 39-50. Импортираните документи в библиотеката могат да имат различни формати, но те се конвертират във вътрешен формат за системата (Greenstone Archive Format). Всеки документ получава автоматично уникален идентификатор OId (Object Identifier).

При добавяне на файлове в библиотеката Greenstone извлича различна метаинформация от тях. Например от HTML-документ може автоматично да се получи заглавие, автор и други описателни данни. Те се използват за съставяне на библиографско описание и за индекси за търсене (контекстно търсене) като допълнение към пълнотекстовото търсене. Ако е невъзможно автоматично извличане на метаданни, то описанието са прави от библиотекар. Форматът на метаданните е съвместим с международния стандарт Dublin Core.

Търсене и навигиране

Greenstone предоставя възможност за осъществяване на пълнотекстово търсене, което може да бъде проведено по структурни части от документа и търсене по елементи от описанието (автор, заглавие и др.). Характерно за този програмен продукт е, че всяка колекция може да има различни критерии за търсене, според документите в нея. Могат да се използват различни методи при търсене: булеви оператори, оператори за групиране, оператори за определяне на разстояние между търсените думи. Преглеждането на документите е по автор, предмет и организация.

Greenstone използва стандарт Unicode за кодировка на символи (ISO/IEC 10646), което позволява документите и потребителският интерфейс да имат различни езикови версии.

Широкото разпространение на програмния продукт Greenstone в цял свят е свързано с мисията и дейността на нейния патрон ЮНЕСКО – да насърчава сътрудничеството между нациите в областта на образованието, науката и културата. Софтуерът е така създаден, че

да се използва и без достъп до интернет. Софтуерът е изтеглен от официалния сайт* от представители на 170 страни⁷⁶. Относно използването му в България не бе открита никаква информация.

Invenio

Макар да не отговаря на критерия за световно разпространение, ще представим накратко програмен пакет за електронна библиотека Invenio, познат в нашата страна. Invenio е използван за създаването на SUDigital – Дигитална библиотека СУ „Св.Климент Охридски”*. Тя съдържа документи, като някои от тях са достъпни след потребителска регистрация. Записите на документите могат да бъдат изведени в различен изходен формат Dublin Core, MARC, MARCXML, BibTeX и др. Програмният пакет използва MARC 21 като основен библиографски формат⁷⁷.

Invenio е пакет с отворен код, който осигурява инструментите за управление на цифрови документи в институционално хранилище. Invenio отговаря на протокола за събиране на метаданни OAI-PMH. Софтуерът обикновено се използва за хранилища с отворен достъп за научно цифрово съдържание и за създаване и поддържане на електронна библиотека. Пакетът е разработен от CERN под названието CDSware, а след 2006 г. се поддържа под името Invenio.

Софтуерният продукт Invenio се използва основно в Швейцария и в някои от скандинавските страни. През последните години се разпространява все по-интензивно. Invenio е избран за създаване на електронни библиотеки на всички национални университети в регионалната икономическа общност UEMOA (Западна Африка), която включва държавите: Бенин, Буркина Фасо, Кот д'Ивоар, Гвинея Бисау, Мали, Нигер, Сенегал, Того⁷⁸.

Доставчикът на услуги TIND Technologies – официален представител на CERN, със седалище в Норвегия предлага Invenio чрез модела „софтуер като услуга”.

* <http://www.greenstone.org/>

⁷⁶ **Factsheet**. In: *Greenstone digital library software* [online]. [Viewed 24.05.2018]. Available from: <http://www.greenstone.org/factsheet>

* <http://lib.sudigital.org/?ln=bg>

⁷⁷ **Invenio**. In: *Wikipedia* [online]. [Viewed 24.05.2018]. Available from: <https://en.wikipedia.org/wiki/Invenio>

⁷⁸ A **virtual** library for 8 UEMOA countries under discussion in Dakar. In: *UNESCO* [online]. UNESCO Office in Dakar, 14 April 2014. [Viewed 24.05.2018]. Available from: http://www.unesco.org/new/en/dakar/about-this-office/single-view/news/a_virtual_library_for_8_uemoa_countries_under_discussion_in/

Изводи:

Прегледът на информационните източници и анализът на практиката в областта на изграждането на електронни библиотеки позволяват да направим следните изводи: 1.

1. Електронната библиотека се определя с базисните характеристики на традиционната библиотека: съхраняване и предоставяне на потребителите на достъп до документи и метаданни за тях, гарантиране на качество и политика на обслужване. Наред с това електронната библиотека предполага някои предимства: едновременен достъп на множество потребители до един и същи документ, предоставяне на нови форми за работа с документите, бързо обновяване на информацията, достъп навсякъде и по всяко време, развити средства и методи за търсене.

2. Електронната библиотека се създава и развива като важна информационна система, предназначена за надеждно съхраняване и ефективно използване на разнообразни колекции от документи (текст, видео, аудио и др.) с унифициран подход към тяхната организация и достъп. Тя има определена архитектура и програмна среда, организация (или лице), която извършва подбор, съхранение и т.н. и гарантира качество на информацията и услугите, политика за използване на електронната библиотека и потребители. Основните функции и задачи на електронната библиотека са интеграцията на информационни обекти (документи и метаданни за тях), предоставянето на достъп, и осигуряване на ефективни средства и методи за търсене и навигиране.

3. Изискванията към електронните библиотеки на съвременния етап на развитие на информационните и комуникационните технологии са за широко и многократно използване на съдържание и за взаимодействие - обмен и интеграция на данни с други информационни системи и/или електронни библиотеки. Това се постига чрез спазването на утвърдени спецификации, стандарти и модели за представяне на информацията. Важна цел е постигането на оперативна съвместимост между електронните библиотеки, особено на семантична оперативна съвместимост. Тя е от съществено значение за осигуряване на възможности за формулиране и изпълнение на сложни информационни запитвания; подобряване точността на търсене; интегриране на хетерогенни информационни ресурси.

Глава 2. РАЗВИТИЕ НА ЕЛЕКТРОННАТА БИБЛИОТЕКА

Днес съществуват сериозни обстоятелства, които затрудняват търсенето и обмена на информация в интернет като метаинформационна среда, където обикновено функционират електронните библиотеки. Някои от тях са свързани с количествени характеристики като увеличаването обема на информацията и разширяване на обхвата, включването на все повече области на човешката дейност в глобалната мрежа. Като основен проблем се отчита нееднозначното възприемане смисъла на информацията или знанията от различните участници в информационния процес: създател – информационна система – получател. Това е предопределено от възможностите на самите информационни системи⁷⁹. В частност, голяма част от документите в уеб представляват неструктурирани или полуструктурирани данни, представени на естествен език. Информацията в тях е неразбираема за машините (компютрите). Търсещата система не „разбира” смисъла на запитването и информацията в документите, а информационно търсене се провежда по срещане на дума/и в предварително индексирани от машини документи. Такова търсене днес се определя като неефективно (налага допълнително филтриране на резултатите) по две съществени причини: несъответствие между контекста на запитването и контекста на резултатите; ключовите думи в запитването не отговарят на използваните думи в документа⁸⁰.

В следствие на тези обективни обстоятелства можем да заключим, че е необходим способ, чрез който данните, информацията и знанията да бъдат структурирани и описани, така че не само получателят (човек), но и компютърът (посредством програмен агент) да бъде способен да „разбира” контекста (смисъла). „Разбира” в този случай означава да обработва данни и информация като използва правила на някакъв логически език, а също и да бъде способен да изведе знания от тях. В днешно време, съобразно очакванията на потребителите, информационните системи се развиват в системи, основани на знания. Това означава, че могат да взаимодействат на естествен език с потребителя, да разбират и използват знания, да правят логически изводи, да предлагат по-ефективно търсене и извличане на информация⁸¹ и др. Тези усъвършенствани информационни системи се

⁷⁹ Тузовский, А. Ф., Чириков, С. В., Ямпольский, В. З. *Системы управления знаниями (методы и технологии)*. Томск: Изд-во НТЛ, 2005, с. 75.

⁸⁰ Пак там, с. 76.

⁸¹ Атанасова, Ирена. *Създаване на експертни системи* [онлайн]. Благоевград: Унив. изд. „Неофит Рилски”, 2018, с. 26–27. [Viewed 20.06.2018].PDF, 4MB. Available from: https://www.researchgate.net/profile/Irena_Atanasova2/publication/323341660_Szdavane_na_ekspertni_sistemi_Expert_systems_development/links/5a8eb5e50f7e9ba42966fe0d/Szdavane-na-ekspertni-sistemi-Expert-systems-development.pdf

създават на базата на най-новите достижения на науките за знанията – изкуствен интелект, компютърна лингвистика и организация на знанията.

Основно понятие в тези науки е „знание”. Знанието се изяснява чрез други основни категории: данни и информация. Данни – информация – знания са свързани понятия, които според степента на абстракция на представата на човека за света образуват пирамида. В основата на пирамидата са данните, следвани от информацията, а знанията са на върха⁸²:

Данните са факти за обект, събитие, процес, получени чрез наблюдение, експеримент или изчисление. Те трябва да бъдат представени във форма, позволяваща съхранението, предаването и обработката им. В „стойността” на данните не се съдържа в явен вид тяхното значение. Данни например е „36 градуса”, но сами по себе си те не могат да бъдат използвани, тъй като не е ясен техният смисъл. Данните могат да бъдат вход на система за обработка на информация, в чиито алгоритми неявно е заложена семантиката им. С оглед на обработката на данните, те се класифицират според структурираното им представяне на:

– структурирани данни – данни, които са групирани в редове и колони в (релационна) база от данни. Така разположени, данните позволяват да се получат отговори на заявки (запитвания) чрез средствата за манипулиране с базата от данни;

– полуструктурирани данни – форма на структурирани данни, при която данните съдържат етикети (метаданни), които уточняват семантиката на данните. Метаданните позволяват данните да бъдат представени в структурирана и организирана форма;

– неструктурирани данни – данни във формат, който не може да бъде представен чрез релационна схема или в самоописващи се структури. Такива са например текст, видео, изображения и др.

Информацията е набор от данни, интерпретирани в определен контекст и предназначени за някаква цел. Следователно информацията се състои от данни и метаданни, описващи данните и тяхното предназначение. Метаданни са данни за данните. Метаданните включват обобщени описания и категоризация на данните.

Знанието е информация не само за смисъла на отделните понятия, но и за връзките и зависимостите между тях (пространствено-времеви, причинно-следствени и т.н.). Знанието е информация, която е обработена, организирана, систематизирана, или обобщена така, че да повиши възприемането, осмислянето или разбирането. Знанията могат да бъдат

⁸² **Bergeron**, Bryan. Essentials of Knowledge Management [online]. Hoboken, New Jersey : John Wiley & Sons, Inc., 2003, p. 10. [Viewed 20.06.2018]. Available from: http://www.cos.ufrj.br/~jano/LinkedDocuments/_papers/aula06/Wiley%20-20Essentials%20of%20Knowledge%20Management.pdf

формално представени и извлечени и от софтуерни приложения. Знанията преминават през различни процеси:

- Възникването на знания е в резултат на установени потребности от нови знания, напр. при различни дейности и събития в живота на отделния човек, производствените практики, научната дейност и т.н., когато се окаже, че наличните знания не са в съответствие с потребностите от знания;

- Извличането на знания е сложен и многообразен процес. Той може да бъде осъществен чрез търсене на знания в книги, списания и др. документални източници, чрез комуникация с експерти, които притежават нужните знания, чрез провеждане на научни изследвания, целящи достигане до нови знания. Поради големите масиви от знания, днес е от особена значимост възможността за автоматично намиране и извличане на знания;

- Кодирането на знания е представянето на знанията в подходяща за използване форма. За хората удобни са текстовите документи, докато за софтуерните агенти знанията трябва да бъдат представени на базата на някои от разработените за целта формализми в бази от данни, бази от знания, системи за организация на знанията като класификации, тезауруси, онтологии и т.н.;

- Съхраняването на знанията се осъществява в документи, хранилища на данни и метаданни, електронни библиотеки, бази от данни и прочие. Начинът на съхранение зависи от начина на представяне на знанията и определя тяхното разпространение и използване;

- Разпространението на знанията може да е свободно или основано на лицензионни условия. За да бъдат разпространявани, знанията трябва да бъдат кодирани и съхранени по подходящ начин. За разпространението на знания е важно тяхното дълготрайно съхраняване, структурирано описание и методите на търсене;

- Използването на знанията зависи не само от тяхната наличност и качество. Важни са и начините им на представяне, възможността потребителите да ги използват (хора и машини). Затова знанията трябва да бъдат представени във вид удобен както за човек, така и за машинна обработка и извличане, за да се използват пълноценно наличните знания;

- Формализацията на знанията се изразява в представянето на понятията от дадена предметна област чрез избран формализъм за представяне на знания. Формализмите за представяне на знания са от изключително важно значение за възможността софтуерни агенти да обработват и използват знания. Формализмите имат два аспекта - синтактичен и логически. Синтактичният аспект определя начина на представяне на знанията (използваните формализми и техния синтаксис), а логическият е свързан с начините за извеждане на нови знания.

В процеса на използването на знания може да се установяват недостиг, непълноти, неточности, неадекватност на знанията с практиката и така се стига до потребността от нови знания и затваряне на цикъла от процеси.

2.1. Науки за знанието, повлияли развитието на електронната библиотека

Науката е форма на човешка дейност за получаване на нови знания за природата, обективния свят и човешкото общество. Някои науки или научни области имат за обект на изследване самото знание. За нас представляват интерес такива съвременни науки, които разглеждат преди всичко формалната организация на знанията във връзка с използването им от хора и машини. Съвременните информационни системи, които обработват информация и знания, използват подходи, методи, технологии, стандарти и средства на различни научни направления: изкуствен интелект, компютърна лингвистика, организация на знанията. Те ще бъдат разгледани накратко с оглед приноса им за развитието на информационните системи в системи, основани на знания, и по-специално за развитието на електронните библиотеки.

2.1.1. Изкуствен интелект

Науката изкуствен интелект е натрупала значителен опит (от 50-те г. на миналия век насам) в моделирането на човешкия интелект в компютърни системи и в създаването на формализми за представяне и използване на знания от компютър. Постиженията на изкуствения интелект – алгоритми и формализми, допринасят за развитието на традиционните информационни системи, трансформирайки ги в системи, основани на знания.

В науката изкуствен интелект са разработени два основни типа формализми според вида на знанията - декларативни и процедурни*. Декларативните формализми служат за представянето на знанията във формат, който позволява лесен и естествен отговор на въпроса „какво?“. Процедурните формализми служат за използването на знанията. При тях знанията се съдържат в процедурите на някаква програма и са представени като отговор на въпроса „как?“. Тъй като процедурните формализми имат приложение при управление на работи, управление на процеси и др.под., където преобладават алгоритмичните знания, на тях няма да бъде акцентирано.

* Декларативни (предметни) знания – знания за обекти и свойства от предметната област; процедурни знания – знания на начините за решаване на задачи в предметната област, както и различни инструкции, методики, включително програмите, които се изпълняват в процеса на решаването на определени задачи.

За нас представлява интерес аспекта на представянето и обработката на знания – модели за формално и формализирано представяне на знания и методи за тяхната обработка. В системите, основани на знания се използват математическа логика, продукционни правила, семантични мрежи и фреймове като основни декларативни формализми. Те ще бъдат разгледани накратко:

– Математическата логика е един от най-рано създадените формализми за представяне и използване на знания. Всяка логическа система е формален език, на който могат да се изразят различни твърдения (знания) и да се извеждат нови знания с помощта на съответни правила за извод. Най-съществени и специфични характеристики на логическите формализми са ясната им семантика и голямата им изразителна сила. Те представят знанията във вид на формули, които се състоят от константи, променливи, функции, предикати, логически връзки и квантори. С тази особеност обаче са свързани и някои недостатъци: прекомерно ниво на формализиране, слаба нагледност, ниска производителност при използване в системите с изкуствен интелект (необходимост от голяма памет)⁸³ и др. Затова логическите модели рядко се използват самостоятелно.

– Продукционни правила е вед формализъм, основан на правила с причинно-следствен характер. Негова основна характеристика е представянето (декомпозирането) на знанията на малки групи от типа „ако-то“: ако условие \rightarrow то действие, ако причина \rightarrow то следствие, ако ситуация \rightarrow то решение. Продукционният модел е подходящ за системи, които решават неголеми задачи. При увеличаване обема на знанията тяхната ефективност спада⁸⁴.

Досега разгледаните формализми математическа логика и продукционен модел се характеризират с това, че знанията се представят посредством правила. Друга група, групата на т.нар. „структурни формализми“ се основава на резултати от изучаването на дълговременната памет на човека и по специално на следните изводи:

- а) Знанията се пазят в мозъка на човека във вид на отделни, неголеми единици, свързани помежду си с множество връзки (асоциации). На това заключение се основава моделът за представяне на знания – „семантична мрежа“;
- б) Човешкият мозък съдържа обособени „порции“ от знания, свързани с дадено централно понятие. Това е основа за създаване на фреймовете.

⁸³ Нишева, Мария М., Шишков, Димитър П. *Изкуствен интелект*. Добрич: Интеграл, 1995, с. 12.

⁸⁴ Лойко, В. И., ред. *Информационные системы и технологии в экономике* [онлайн]. Москва: Финансы и статистика, 2005. [Прегледан 12.03.2018]. PDF, 19.1 МВ. Достъпно от: <http://lib.maupfib.kg/wp-content/uploads/2015/12/end/akademy/03%20inf%20sis%20i%20teh%20v%20econ/inf%20sis%20i%20teh%20v%20econ%20loiko.pdf>

Семантичната мрежа представя знанията с помощта на мрежа от възли (върхове) и дъги, които ги свързват, т.е. по своята структура представлява граф. С възлите се означават обектите (понятия, събития, процеси, явления) от съответна предметна област, а с дъгите - връзките (отношенията, релациите) между тях. Изводите се правят при проследяване на връзките със съответните понятия, например:



Коко е папагал. Папагалът е птица. Птицата има крила.

“Коко”, “папагал”, “птица”, “крила” са понятия; “е”, “е”, “има” са отношения между понятията.

Семантичните мрежи могат да отразяват различни отношения: обект – клас и клас – суперклас („IS A”); част–цяло; функционални връзки (определят се от глаголи, напр. „направен от ...”, „влие на ...”); количествени отношения (много, малко, равно); пространствени отношения (далече от, близо до...); времеви отношения (по-рано, по-късно, по време на ...); атрибутивни връзки (има свойство, има значение); логически връзки (И, ИЛИ, НЕ) и др.

Основни характеристики на семантичните мрежи са, че описват знанията на естествен език и ги представят в относително еднородна структура. Като предимства на този формализъм се сочат: гъвкавост, постигана благодарение на отношенията „йерархичност” и „асоциативност”, нагледност – отразяване на понятията и връзките между тях чрез графични нотации; лесна четимост и разбиране на знанията; висока степен на структурализация на знанията⁸⁵.

Фрейм е метод за представяне на знания чрез абстрактен образ (рамка), т.е. минимално, най-общо описание на някаква същност (обект, явление, събитие, или процес). Авторът на теорията за фреймовете Марвин Мински взема за отправна точка факта, че човек, опитвайки се да опознае нова ситуация, използва от своя опит някаква структура от данни (образ), наречен от него „фрейм”. Този фрейм с малки изменения може да се използва за разбиране на широк кръг от ситуации, явления или процеси⁸⁶. Мински дава следния пример: преди да влезе в стая, човек вече има в съзнанието си образа на понятието „стая” – стени, таван, под и т.н., когато влезе в стаята, той само допълва и конкретизира своята

⁸⁵ Нишева, Мария М., Шишков, Димитър П. Цит. съч., с. 11.

⁸⁶ Minsky, Marvin. *A framework for representing knowledge* [online]. Boston, Massachusetts: MIT AI Laboratory Memo 306, June, 1974. [Viewed 20.06.2018]. Available from: <http://web.media.mit.edu/~minsky/papers/Frames/frames.html>

представа за „стая“ с особеностите на конкретната стая. Идеята за фреймовете е аналог на създаването и използването на представи, аналогии и натрупан (предишен) опит в човешкото мислене.

От разгледаните формализми фреймовете предлагат най-висока степен на структуриране на знанията. Освен това фреймовете се характеризират с добра нагледност и модулност. Като недостатък, както и при семантичните мрежи се сочи отсъствието на универсалност по отношение на процедурите по обработката, което води до неефективно използване на изчислителната техника по отношение на памет и бързо действие⁸⁷.

Опит да се преодолеят някои недостатъци е съчетаването на различни модели: фреймове и продукционни правила; семантични мрежи и логически модели; семантични мрежи и продукционни правила.

Напредъкът в областта на изкуствения интелект позволява на системите да решават задачи за търсене, да извършват гласово разпознаване, автоматичен превод на текст, синтез на реч, превод на реч на различни езици в реално време⁸⁸ и др.

2.1.2. Компютърна лингвистика

Според Марк Стийдман „човешкото знание се изразява чрез език, така че компютърната лингвистика е много важна“⁸⁹.

Теоретико-приложното поле на компютърната лингвистика се формира като научно-изследователска област през 50-те години на миналия век с усилията на американски изследователи да използват компютрите за автоматичен превод на текстове от чужди езици, най-вече от руски на английски език⁹⁰. Практическите приложения на машинния превод не покриват първоначалните очаквания на изследователите, нито за лесна приложимост, нито за точност. Автоматизираната обработка на естествен език се оказва много по-сложна задача. През 60–70-те г. компютърната лингвистика се развива в посока разработване на теоретично обосновани алгоритми за интелигентна обработка на езикови данни. Терминът

⁸⁷ Нишева, Мария М., Шишков, Димитър П. Цит. съч., с. 11.

⁸⁸ Атанасова, Ирена. *Създаване на експертни системи* [онлайн]. Благоевград: Унив. изд. „Неофит Рилски“, 2018, с. 26–27. [Viewed 20.06.2018].PDF, 4MB. Available from: https://www.researchgate.net/profile/Irena_Atanasova2/publication/323341660_Szdavane_na_ekspertni_sistemi_Expert_systems_development/links/5a8eb5e50f7e9ba42966fe0d/Szdavane-na-ekspertni-sistemi-Expert-systems-development.pdf

⁸⁹ Цит. по: **Computational Linguistics**. In: *Stanford Encyclopedia of Philosophy* [online]. [Viewed 28.02.2018]. Available from: <https://plato.stanford.edu/entries/computational-linguistics/>

⁹⁰ **Hutchins**, John. Retrospect and prospect in computer-based translation. In: *Proceedings of MT Summit VII* [online]. 1999, pp. 30–44. [Viewed 20.06.2018]. Available from: <https://pdfs.semanticscholar.org/bb87/8ee8c564faf45195adcd0c59a9348e336ddb.pdf>

„компютърна лингвистика” е въведен от Дейвид Хайс*. През 80-те години до 1995 г. компютърната лингвистика се основава на статистически подходи, базирани на езикови корпуси⁹¹.

Постиженията на съвременната компютърна лингвистика често са резултат от комбинирането на методи и подходи от областта на компютърните науки и програмирането, математиката, статистиката, лингвистиката, изкуствения интелект и др. Практическите задачи на компютърната лингвистика са широки и разнообразни. Тя се занимава както със „създаване на езикови технологии и изчислителни модели, така и с тяхното приложение в системи, които влияят на качеството и ефективността на общуването между хората и машините за проверка на правописа, автоматичен превод, търсене и извличане на информация, трансформиране на писмен текст в реч и обратно, разпознаване на текст от изображения и много други”⁹².

Изключително важна задача на компютърната лингвистика е моделирането на семантиката в езика и построяването на лексикални семантични ресурси (електронни речници, лексикални онтологии и семантични мрежи). Семантични ресурси и инструменти се използват при решаването на много задачи, свързани с обработката на естествен език: резюмиране на текст, категоризиране на текст, търсене и извличане на информация от текст, отговаряне на въпроси, извличане на знания, извличане на факти, разрешаване на многозначност, перифразиране на текст и идентификация на перифразиран текст, генериране на текст, машинен превод, семантичен анализ и други.

Интересът към изграждането и използването на семантични ресурси и инструменти за извличане и анализиране на семантични връзки в уеб се доказва със значителния брой научни публикации, свързани с тях⁹³. Например само публикациите, свързани с WordNet –

* Дейвид Хайс (*David Hays, 1928–1995*) играе важна роля в професионалното налагане на компютърната лингвистика. Той е сред основателите на авторитетните в тази област организации – Асоциацията по компютърна лингвистика (*Association for Computational Linguistics, 1962–*) и Международния комитет по компютърна лингвистика (*International Committee on Computational Linguistics, 1965–*). Той е първият редактор (1974–1978 г.) на специализираното научно списание „*Computational Linguistics*”. Практическата дейност на Хайс за корпорацията RAND (САЩ) като ръководител на проект за съставянето на корпус от един милион думи от руски текстове, го прави пионер в направлението, сега известно като корпусна лингвистика.

⁹¹ Повече за развитието на компютърната лингвистика в: **Computational Linguistics** [online]. [Viewed 20.06.2018]. Available from: <https://plato.stanford.edu/entries/computational-linguistics/>

⁹² Коева, Светла. Предговор. В: *Езикови ресурси и технологии за български език*. София: Акад. изд. „Проф. Марин Дринов”, 2014, с. 7.

⁹³ Након, Светлин. Автоматично извличане на фалшиви приятели от паралелен двуезичен корпус [онлайн]: Автореферат на дисертационен труд. София, 2009, с. 3. [Прегледан 28.02.2018]. Достъпно от: <https://www.nakov.com/wp-content/uploads/2010/05/nakov-phd-avtoreferat-false-friends.pdf>

един от основните семантични ресурси за английски език – към февруари 2018 г. са около 97 800 резултата според Google Scholar.

WordNet е лексикална онтология, изградена във вид на семантична мрежа. Във възлите на графа на семантичната мрежа WordNet се разполагат дума/думи, (могат да бъдат прости думи, сложни думи или съставни думи), наречени литерали, които назовават понятие и образуват синонимно множество (синсет). Дъгите, които свързват възлите, изразяват отношения (релации) между синонимните множества. Семантичните релации между литералите в WordNet могат да бъдат: синонимия, антонимия, хиперонимия/хипонимия, меронимия/холонимия⁹⁴.

Например* литералът ВЛАК е с релация *синонимия*, т.е. в един синсет с КОМПОЗИЦИЯ, ВЛАКОВА КОМПОЗИЦИЯ, което се описва с тълковната дефиниция „вид обществено транспортно средство, представляващо скачени един за друг вагони, теглени от локомотив по релсов път и предназначено за превозване на пътници, поща и товари“. С релацията “хипероним” (род-вид) това синонимно множество е свързано с по-общото синонимно множество ОБЩЕСТВЕН ТРАНСПОРТ – „организирана система за наземен, подземен (рядко воден или въздушен) редовен превоз на пътници, товари и поща по определени вътрешноселищни и междуселищни маршрути, осигурявано от държавата или от частни фирми“. С релацията “хипонимия” (вид-род) е свързано с няколко по специфични синонимни множества: ВЛАК, КОЙТО ПРЕВОЗВА ПЪТНИЦИ ДО ПРИСТАНИЩЕ; ТОВАРЕН ВЛАК и ТОВАРНА КОМПОЗИЦИЯ; БОЛНИЧЕН ВЛАК, САНИТАРЕН ВЛАК; ПОЩЕНСКИ ВЛАК; ПЪТНИЧЕСКИ ВЛАК; ПОДЗЕМЕН ВЛАК, ВЛАК НА МЕТРОТО. С релацията “меронимия” (цяло-част) синонимното множество, в което е включено ВЛАК е свързано със синсети, които изразяват неговите съставни части ВАГОН и ЛОКОМОТИВ. Респективно синсет, който представлява част, е с релацията “холонимия”. В случая избрания пример не илюстрира релация “антонимия” (противоположност), която обикновено характеризира прилагателните имена, напр. синсет ВИСОК има релация “антонимия” със синсета НИСЪК. Семантичната мрежа може да отрази и междуезикови релации. Например с междуезикова релация за еквивалентност синсетът ВЛАК, КОМПОЗИЦИЯ, ВЛАКОВА КОМПОЗИЦИЯ е

⁹⁴ Повече за WordNet виж: **Miller**, George A., **Beckwith**, Richard, **Fellbaum**, Christiane, **Gross**, Derek, **Miller**, Katherine. *Introduction to WordNet: An On-line Lexical Database* [online]. Revised August 1993. [Viewed 28.02.2018]. Available from: <http://wordnetcode.princeton.edu/5papers.pdf>

* Примерът за литерали, синонимични множества, дефиниции и релациите е от Българския WordNet достъпен от: <http://dcl.bas.bg/bulnet/>

свързан със синонимното множество на английски език TRAIN, RAILROAD TRAIN, възможно и с други езици.

Началото на WordNet е положено през 1985 г. в Лабораторията за когнитивни науки на Принстънския университет от група учени психолози и лингвисти с ръководител проф. Джордж Милър⁹⁵. Първоначалната цел е да се изгради лексикална база от данни като модел на човешката памет. През 1995 г. WordNet е предоставен за свободно ползване в интернет, за провеждане на различни научни изследвания и прилагане в компютърни приложения за автоматична обработка на текст. Днес версия 3.0 на Принстънският WordNet съдържа 117 659 синонимни множества, разпределени в четири части на речта: съществителни, прилагателни, глаголи и наречия, в които са включени 155 287 литерала, а еднозначните отношения литерал-значение са 206 941⁹⁶. След Принстънския WordNet се разработват около 80 подобни електронни бази от данни за повече от 40 езика⁹⁷. Отличителен знак за голямото му влияние е, че името на Принстънската семантична мрежа – WordNet, става нарицателно име, означаващо всички организирани по подобен начин езикови ресурси. Идеята се разгръща в проекта EuroWordNet – създаване на многоезична база от данни, включваща няколко европейски езика – английски, естонски, испански, италиански, немски, френски, холандски и чешки* (1996–1999 г.). Проектът въвежда междуезиков индекс (Inter-Lingual Index, ILI), чрез който всяко отделно синонимно множество се свързва със съответното множество в английския език, а чрез английски – и с всички останали езици. Синонимните множества от междуезиковия индекс играят ролята на езиково независими понятия (concepts), които позволяват едноезичните лексикално-семантични мрежи да се превърнат в многоезична лексикална онтология. След завършването на проекта много други институти и изследователски групи могат да разработват съвместими „урднети“ на други езици, като използват спецификацията на EuroWordNet. Такива се разработват за шведски, норвежки, датски, гръцки, румънски, литовски, руски, български, словенски и др.

По проект на Българската академия на науките (БАН) BalkaNet** (2001–2004) са създадени синхронизирани семантични бази от данни за следните балкански езици –

⁹⁵ Miller, George A., Beckwith, Richard, Fellbaum, Christiane, Gross, Derek, Miller, Katherine. Op. cit.

⁹⁶ wnstats (7WN). In: *WordNet: A Lexical Database for English* [online]. [Viewed 20.06.2018]. Available from: <https://wordnet.princeton.edu/documentation/wnstats7wn>.

⁹⁷ Wordnets in the World. In: *Global WordNet Association* [online]. [Viewed 20.06.2018]. Available from: <http://globalwordnet.org/wordnets-in-the-world/>

* <http://projects.ilic.uva.nl/EuroWordNet/>

** <http://www.dblab.upatras.gr/balkanet/index.htm>

български, гръцки, румънски, сръбски, турски, и разширяване със съществуващата чешка лексикално-семантична мрежа. Основната цел на проекта е както изграждането на електронни семантични мрежи за шестте изброени езика, така и демонстрирането на тяхното приложение в областта на автоматичната обработка на естествените езици⁹⁸. Една от важните задачи при изграждането на лексикално-семантичната многоезична мрежа на балканските езици е запазването на връзката с английския WordNet (а следователно и с европейските лексикално-семантични мрежи), за да могат да се осъществяват автоматични преходи между различните езици. Изискването за съответствие налага постоянна координация с промените и подобренията, правени в Принстънския WordNet⁹⁹.

Българският WordNet* е съставен и развит в рамките на национално финансираните проекти „BulNet – лексикално-семантична мрежа на българския език” (2005–2010) и „Електронни езикови ресурси и програми за тяхната обработка” (2011–2013) на Секция по компютърна лингвистика към БАН¹⁰⁰. BulNet представлява електронен многоезичен ресурс, съдържащ синоними с тълковни дефиниции и още седемнадесет, предимно семантични релации, като антоними, мероними, каузация, област на употреба и др.¹⁰¹ Към 2014 г. българската база от данни съдържа 50 226 синонимни множества, разпределени в частите на речта, които съответстват на съществителните, глаголите, прилагателните и наречията в Принстънския WordNet, като освен тях са включени и местоимения, предлози, съюзи, частици и междуметия¹⁰². Екипът от българска страна се състои основно от експерти от Секцията по компютърна лингвистика в Института за български език към БАН – компютърни лингвисти, математици, логици и програмисти.

Базата от данни BulNet е интегрирана в мрежата на балканските езици BalkaNet и на европейските езици EuroWordNet чрез междуезиковия индекс, който означава еднозначно съответните синонимни множества в отделните езици. BulNet предлага решения на семантично равнище: възможност за избор на синоними, възможност за справка за семантичните връзки на дадена дума по отношение на системата от останали думи в езика

⁹⁸ **BalkaNet** – Design and Development of a Multilingual Balkan WordNet. In: *Balkanet Word* [online]. [Viewed 20.06.2018]. Available from: <http://www.dblab.upatras.gr/balkanet/index.htm>

⁹⁹ Ibid.

* <http://dcl.bas.bg/bulnet/>

¹⁰⁰ **Wordnet**. In: *Секция по компютърна лингвистика* [онлайн]. [Прегледан 20.06.2018]. Достъпно от: <http://dcl.bas.bg/resursi/wordnet/>

¹⁰¹ **Семантичната** мрежа на балканските езици е част от световната WordNet. В: *Computer world* [онлайн]. 2 март 2006. [Прегледан 28.02.2018]. Достъпно от:

http://computerworld.bg/11835_semantichnata_mrezha_na_balkanskite_ezici_e_chast_ot_svetovната_wordnet

¹⁰² **Коева**, Светла. WordNet и BulNet. В: *Езикови ресурси и технологии за български език*. София: Акад. изд. „Проф. Марин Дринов”, 2014, с. 154–173.

(антоними, отношения род-вид, цяло-част и др.), възможност за справка с тълковното значение на думата и паралелни предложения за съответната дума на английски, немски, френски, испански, италиански, холандски, чешки, естонски, гръцки, румънски, турски и сръбски¹⁰³.

Съвременните езикови технологии, създавани в резултат на постиженията на компютърната лингвистика, позволяват изграждането на справочници за езика, каквито не е имало досега, които се използват за автоматичната обработка на естествен език с цел извличане на информация и знания. По-важните приложения на многоезичната лексикално-семантична мрежа са напр. отстраняване на семантична многозначност, автоматично отговаряне на въпроси, автоматично резюмиране на документи, автоматична категоризация на документи, автоматичен превод, автоматично извличане на информация и др. За нашето изследване представлява интерес приложението на лексикалните семантични мрежи за търсенето и извличането на информация. Представените „урднети“ са организирани на основата на понятия и техните значения, което ги прави приложими за концептуално индексирание и търсене. В тази насока се провеждат следните изследвания: а) търсене на документи не по отделна дума, а по значението (тълковната дефиниция), б) разширяване на информационното запитване със синоними или други думи, свързани с тези от запитването, чрез проследяване на релации.

2.1.3. Организация на знанието

За да бъде знанието използваемо от хора и технологични агенти, то трябва да бъде организирано по някакъв начин. Организирането на знанията става все по-важно предвид непрекъснато нарастващото научно и културното производство в течение на годините¹⁰⁴.

Организация на знанието (Knowledge organization) е област на научни изследвания, образование и практическа дейност, която е тясно свързана с библиотечно-информационните науки. Тя се институционализира като:

– академична специалност или дисциплина (например специалност Организация на знанието в Royal Department of Continuing Education and Consultancy към School of Library and Information Science in Copenhagen (Дания) и University College in Borås (Швеция),

¹⁰³ Семантичната мрежа на балканските езици...

¹⁰⁴ Soergel, Dagobert. Digital Libraries and Knowledge Organization [online]. In: Kruk, Sebastian Ryszard, McDaniel, Bill, eds. *Semantic Digital Libraries*. Berlin: Springer, p. 9–39. [Viewed 20.06.2018]. Available from: https://link.springer.com/chapter/10.1007/978-3-540-85434-0_2

магистърска програма в University of Washington Information School (САЩ), курс във Faculty of Information на University of Toronto (Канада);

- научноизследователски организации (напр. Международно дружество за организация на знанието (International Society for Knowledge Organization, ISKO);
- специализирана периодика (напр. списанието „Knowledge Organization”);
- международни форуми (напр. конференцията Knowledge Organization – what's the story?, 11th–12th September 2017, London).

Дълголетният опит на библиотеките и останалите „институции на паметта” за представяне, организиране и систематизиране на колекции (от издания, архивни документи или експонати) допринася за формирането на “Организация на знанието” като отделна академична и научноизследователска област. Някои изследвания я разглеждат като поле в рамките на библиотечно-информационните науки. Според други изследователи “Организацията на знанието” е интердисциплинарно поле, много по-широко от Библиотечно-информационните науки¹⁰⁵.

Важен принос за развитието на тази научна област имат изследователи като Чарлс Кътър, Джон Дюи, Ернст Ричардсън, Уилям Б. Сайърс, Хенри Блис, Шиали Рамамрита Ранганатан и др. Те налагат схващането, че класификацията на документите трябва да се основава на организацията на знанията в науката¹⁰⁶. Блис първи използва термина „организация на знанието” в две от своите книги¹⁰⁷. Идеята споделена от Х. Блис и Дж. Дюи, както и от много съвременни изследователи е, че науките отразяват реда на природата, а библиотечно-библиографските класификации трябва да представят знанията така, както са организирани в науката:

Естествен ред → Класификация на науките → Класификация на библиотеките¹⁰⁸.

¹⁰⁵ **Hjørland**, Birger. Fundamentals of Knowledge Organization [online]. In: *Knowledge Organization*, 2003, Vol. 30, Issue 2, p. 87–111. [Viewed 20.06.2018]. PDF, 1MB. Available from: <http://ppggoc.eci.ufmg.br/downloads/bibliografia/Hjorland2003.pdf>

¹⁰⁶ **Hjørland**, Birger. Knowledge organization. *Knowledge Organization*, 2016, Vol. 43, Issue 6, pp. 475–484. Also available in Hjørland, Birger, ed. *ISKO: Encyclopedia of Knowledge Organization* [online]. [Viewed 20.06.2018]. Available from: http://www.isko.org/cyclo/knowledge_organization

¹⁰⁷ **Bliss**, Henry. *The Organization of Knowledge and the System of the Sciences* [online]. New York: H. Holt and Company, 1929; **Bliss**, Henry. *The Organization of Knowledge in Libraries and the Subject-Approach to Books*. New York: Wilson, 1933.

¹⁰⁸ Идеята е подчертана специално от автора на предговора Джон Дюи в монографията на Хенри Блис виж: **Bliss**, Henry. *The Organization of Knowledge and the System of the Sciences*. New York: H. Holt and Company, 1929, p. viii. [Viewed 20.06.2018]. Available from: <https://babel.hathitrust.org/cgi/pt?id=mdp.39015011706697;view=1up;seq=7>

По-късно темата е подета от немската изследователка Ингетрат Далберг. Тя използва термина „организация на знанието“ (нем. Wissensordnung) за означаване на концептуалната и систематична организация на човешкото знание¹⁰⁹. Нейната дейност в тази област е изключително плодотворна. Тя основава специализираното научно периодично издание *Knowledge Organization: International Journal devoted to Concept Theory, Classification, Indexing, and Knowledge Representation*^{*} и Международното общество за организация на знанието (International Society for Knowledge Organization, ISKO).

Още от пионерите в тази научна област се налагат две различни по обхват значения на понятието „организация на знанията“. Според по-широкото разбиране, организацията на знанието е свързана с всички сфери на обществото и културата¹¹⁰. Ричард Уитли смята, че организацията на знанието има два аспекта: а) как се организира знанието в обществото (например разделение по научни области и дисциплини, по професии и т.н., т.е. „социална организация на знанието“) и б) как концептуално знанието се организира в принципи, теории, научни таксономии (например периодичната система в химията или таксономииите на биологичните видове), тоест - „познавателната организация на знанието“. Авторът отбелязва, че има взаимовръзка между тези два аспекта. Този по-широк смисъл на понятието организация на знанията обслужва природните науки, метафизиката и социология на знанието¹¹¹.

Тесният смисъл на „организация на знанието“ е свързан с дейности като описание, индексирание и класификация на документи (или експонати), извършвани в библиотеките, музеите, архивите, библиографските агенции и информационните центрове. Тези дейности в съвременното информационно пространство могат да бъдат извършвани както от професионалисти (библиотекари, архивисти, информационни специалисти), така и от компютърни алгоритми и крайни потребители (например социално тагиране).

¹⁰⁹ **Dahlberg**, Ingetraut. *Grundlagen universaler Wissensordnung: Probleme und Möglichkeiten eines universalen Klassifikationssystems des Wissens*[online]. Pullach: Verlag Dokumentation. [Viewed 20.06.2018]. Available from: https://books.google.bg/books?id=LYnhOLiUPSIC&printsec=frontcover&hl=bg&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false

^{*} в периода 1974-1993 излиза със заглавие *International Classification*.

¹¹⁰ виж например: **Hjørland**, Birger. *Fundamentals of Knowledge Organization...* [online]; **Hjørland**, Birger. *What Is Knowledge Organization (KO)* [online]. In: *Knowledge Organization, 2008*, Vol. 35, Issue 2–3), p. 86–102 [Viewed 28.06.2018]. Available from:

https://www.researchgate.net/publication/277803483_What_is_Knowledge_Organization_KO.

¹¹¹ **Whitley**, Richard R. *The Intellectual and Social Organization of the Sciences* [online]. Oxford: Oxford University Press, 1984 [Viewed 28.06.2018]. Available from:

https://www.researchgate.net/publication/235413219_The_Intellectual_and_Social_Organization_of_the_Sciences

Разбира се „организацията на знанието” като средство за систематизиране на информационни ресурси е тясно свързана с библиотечно-информационните науки. То се разглежда като: а) процесите по организация на знанието (knowledge organization processes, KOPs) – каталогизиране, индексирание, предметизиране, класифициране и б) инструментите и средствата, предназначени за организиране на знания и информация и за улесняване на тяхното управление и търсене¹¹², наречени общо системи за организация на знания.

Направлението на науката “Организация на знанията”, което е посветено на изследването на системите за организация на знанията и тяхното приложение в информационните системи е свързано с предмета на нашето изследване. Част от системите ще бъдат разгледани като информационни езици, изграждащи лингвистичното осигуряване на електронна библиотека.

Нашата цел тук е да разгледаме съществуващите системи за знанията, като ги представим като типово-видово разнообразие и опишем техните специфики.

Самият събирателен термин „системи за организация на знанията” (knowledge organization system) е предложен и въведен на първата среща на Работната група на Networked Knowledge Organization Systems (NKOS Working Group), проведена на конференцията на ACM Digital Libraries 1998 в Питсбърг. Терминът придобива голяма популярност, когато консорциумът W3 започва да го използва в своите разработки (спецификации и стандарти), свързани със семантичния уеб. През 2008 г. консорциумът публикува стандарта “Система за опростена организация на знанията” (Simple Knowledge Organization System, SKOS) за представяне на тезауруси, класификационни схеми, предметни рубрики, таксономии, фолксономии и др. в уеб като свързани данни¹¹³. Стандартът ще бъде разгледан като технология на семантичния уеб в т. 2.2.3 на настоящата глава.

Няма общоприето определение на понятието „системи за организация на знанията”. Ние се придържаме към определението на често цитирания автор Дагоберт Соергел¹¹⁴:

¹¹² Виж например: **Broughton**, Vanda, **Hansson**, Joacim, **Hjørland**, Birger, **Huertas**, María José López. Knowledge Organization [online]. In: *European Curriculum Reflections on Library and Information Science Education*. Copenhagen: Royal School of Information Science, p. 133–148 [Viewed 28.06.2018]. Available from: <http://www.webcitation.org/6q0hjT516> и **Hjørland**, Birger. What Is Knowledge Organization (KO)... [online].

¹¹³ **SKOS** Simple Knowledge Organization System. [online]. Reference W3C Recommendation 18 August 2009. [Viewed 20.06.2018]. Available from: <https://www.w3.org/TR/skos-reference/>; **SKOS** Simple Knowledge Organization System Reference [online]. W3C Working Draft 25 January 2008. [Viewed 20.06.2018]. Available from: <http://www.w3.org/TR/2008/WD-skos-reference-20080125/>

¹¹⁴ **Soergel**, Dagobert. *Knowledge Organization Systems: Overview* [online]. p. 3 [Viewed 20.06.2018]. Available from: <http://www.dsoergel.com/SoergelKOSOverview.pdf>

Системи за организация на знанието обхваща широк кръг от системи, като онтологии, схеми за описание (метаданни), таксономии, класификации, структури на уеб директории, тезауруси, речници, фолксномии и др., които имат различно предназначение. Общото между тях е, че служат за организиране, търсене и откриване на информация и знания за света от хора и машини. Освен това системите позволяват на компютърните програми да правят съждения.

Система за организация на знанието може да се прилага за различни данни за ресурса (автор, идентификатор, предмет, класификационен индекс и др.), вградени в ресурса като етикети (тагове) или отделени от ресурсите на електронната библиотека и използвани като част от механизма за достъп.

Изборът на система/и за организация на знанието, използвана в конкретна електронна библиотека може да бъде взаимствана от практиката на библиотеките (напр. Универсална десетична класификация, Library of Congress Classification Schedule), от друга информационна система (напр. темите Yahoo categories) или могат да бъдат специално създадени.

Тъй като съществуващите инструменти за организация на знанията са най-разнообразни по състав, структура, практическо приложение, цел на създаване и др., изследователите от тази област правят различни опити за тяхната типологизация и класификация¹¹⁵.

Може би една от най-изчерпателните типологизация на системите за организация на знанията е предоставена от Ренато Суза и авторски колектив¹¹⁶ (Фиг. 11).

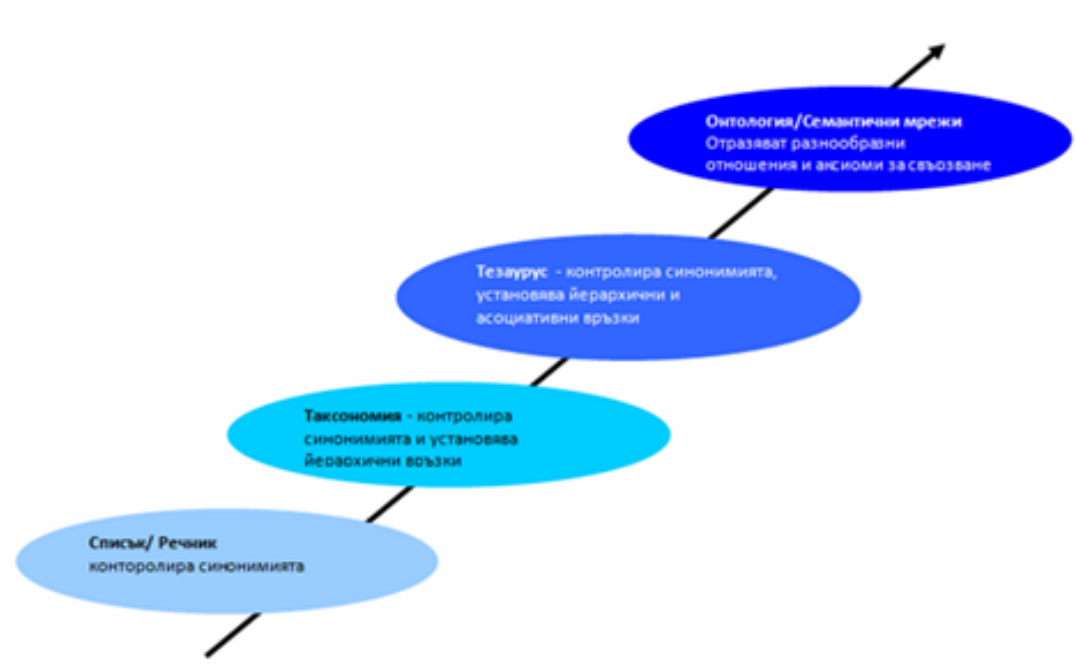
¹¹⁵ виж например: **Bergman**, Michael K. *An Intrepid Guide to Ontologies* [online]. [Viewed 20.06.2018]. Available from: <http://www.mkbergman.com/?p=374>; **Soergel**, Dagobert. Evaluation of Knowledge Organization Systems (KOS): Characteristics for Describing and Evaluating KOS [online]. In: *Paper presented at Classification Crosswalks: Bringing Communities Together. The 4th NKOS Workshop at ACM-IEEE Joint Conference on Digital Libraries (JCDL), June 28th, Roanoke, VA* [Viewed 20.06.2018]. Available from: <http://nkos.slis.kent.edu/2001/SoergelCharacteristicsOfKOS.pdf>; **Tudhope**, Douglas, **Koch**, Traugott, **Heery**, Rachel. *Terminology Services and Technology: JISC State of the Art Review* [online]. Bath: UKLON [Viewed 20.06.2018]. Available from: <http://www.ukoln.ac.uk/terminology/JISC-review2006.html>; **Wright**, Sue Ellen. Typology for KRRs [online]. In: *Networked Knowledge Organization Systems / Services / Structures NKOS*. [Viewed 20.06.2018]. Available from: <http://nkos.slis.kent.edu/2008workshop/SueEllenWright.pdf>

¹¹⁶ **Souza**, Renato Rocha, **Tudhope**, Douglas, **Almeida**, Mauricio B. Towards a Taxonomy of KOS: Dimensions for Classifying Knowledge Organization Systems [online]. In: *Knowledge Organization, Vol. 39, Issue 3*, p. 179–192. [Viewed 20.06.2018]. Available from: http://mba.eci.ufmg.br/downloads/Souza_Tudhope_Almeida_-_KOS_Taxonomy.Submitted.pdf

Според авторите на тази типология всички инструменти или средства, които се използват за организиране на знанията, за информационно търсене, за поддържане управлението на знанията или за представяне на знания чрез терминология, трябва да се разглеждат като системи за организация на знанията. Затова предложената от тях типология включва резюмета, семантични мрежи, информационно-търсещи индекси (IR indexes), но не включва маркиращи езици като HTML и SGML, които представляват формати за тяхното представяне*.

Системите за организация на знанията се различават по своята семантична сила, т.е. възможността да отразяват семантични връзки между термините/понятията. Много изследователи се позовават на идеята за „скала на семантиката”, предложена от А. Блумауер и Т. Пелигрини¹¹⁸ и използват степента на семантично богатство на системите за организация на знания като критерий за сравняване и класифициране (Фиг. 12).

Фигура 12. Скала на семантиката на системите за организация на знанията



За целите на настоящото изследване се фокусираме върху структурата и типичните свойства на системите за организация на знанията, използвани в библиотеките и предлагаме

* Маркиращите езици са включени в типологията на Е. Урайт виж: **Wright**, Sue Ellen. Typology for KRRs [online]. In: *Networked Knowledge Organization Systems/Services/Structures NKOS*. [Viewed 20.06.2018]. Available from: <http://nkos.slis.kent.edu/2008workshop/SueEllenWright.pdf>

¹¹⁸ **Blumauer**, A., **Pellegrini**, T. Semantic Web und semantische Technologien [online]. In: *Semantic Web: Wege zur vernetzten Wissensgesellschaft*. Berlin, Heidelberg: Springer. p.16 [Viewed 20.06.2018]. https://ch.semweb.ch/_/wordpress/wp-content/uploads/2013/08/semantic-webscool-wien.pdf

следната типологизация, взaimствана от изследванията на Ходж¹¹⁹. Тя разделя системите за организация на знанията в 3 категории: списък на термини, класификации или категоризации и системи от термини или понятия с връзките между тях (табл. 4).

Таблица 4. Системи за организация на знанията по Ходж

Категории	Характеристики	Видове
Списъци на термини	Линейни или слабоструктурирани системи; акцент върху списъците на термините (възможно с контролиране на синонимията, т.е. отразяване на еквивалентни връзки)	Нормативни файлове Глосари Речници Географски справочници
Класификации и категоризации	Йерархично структурирани системи; акцент върху създаването на тематични класове (категории) с йерархични връзки между тях.	Предметизационни системи Класификационни системи Таксономии Категоризации Фолксономии
Термини/понятия, и семантични връзки	Сложноструктурирани системи; акцент върху връзките между термините или понятията (еквивалентни, йерархични, асоциативни)	Тезауруси Концептуални карти Тематични карти Семантични мрежи Онтологии

Видове и предназначение на системите за организация на знания:

Списъци на термини:

– Нормативни файлове (Authority files). Нормативните файлове са списъци с термини, които се използват за контролиране на форми на имена или названия на същности като географски наименования, имена на лица или наименования на колективни органи и

¹¹⁹ Hodge, Gail. *Systems of Knowledge Organization for Digital Libraries: Beyond Traditional Authority Files* [online]. Washington, DC: The Digital Library Federation, April 2000, 37 p. [Viewed 20.06.2018]. Available from: <https://www.clir.org/wp-content/uploads/sites/6/pub91.pdf>

др. Обикновено вариантите форми са свързани с предпочетена форма. Този тип система за организация на знанията обикновено не включва задълбочена организация или сложна структура. Представянето може да бъде азбучно или организирано чрез елементарна класификационна схема. Може да се използва семпла йерархия, за да се позволи лесна навигация, особено когато нормативният списък е изключително голям. Примери: *Library of Congress Name Authority File* и *Getty Geographic Authority File*.

– Глосари (Glossaries). Глосарът е списък с термини, обикновено с определения. Термините могат да бъдат от конкретно предметно поле или свързани с определена практическа дейност. Термините се дефинират с оглед използването в точно определена област или приложение и рядко включват вариантни значения. Примери: *Glossary of Environmental Terms*, *Environmental Protection Agency Glossary*.

– Речници (Dictionaries). Речниците са азбучни списъци с думи и техните определения. Речниците са по-общи по тематичен обхват, отколкото глосарите. Те могат също така да предоставят информация за произхода на една дума, варианти (по правопис и морфология) и различни значения. Речникът обикновено представя термините без да прилага явна йерархична структура. Примери: *Oxford Dictionaries*, *Dictionary by Merriam-Webster*.

– Географски справочници (Gazetteers). Географските справочници съдържат информация за географски обекти (географски, демографски и др. статистически данни) и техните наименования. Традиционните географски справочници са публикувани като книги или съставяни като показалци към атласи. Географските обекти могат да бъдат представени азбучно или да бъдат групирани по категории (чрез класификационна схема), напр. градове, реки и т.н. Електронните географски речници съдържат списък с имена на места, заедно с местоположение в географска ширина и дължина (геолокация)¹²⁰, напр. *JRC Fuzzy Gazetteer (на Joint Research Centre of the European Commission*, U.S. Board on Geographic Names***.

¹²⁰ Повече за географските справочници виж: **United Nations Group of Experts on Geographical Names**. Manual for the national standardization of geographical names [online] [Viewed 20.06.2018]. Available from: New York: United Nations, 2006, p. 73-76. [Viewed 30.05.2018]. Available from: https://unstats.un.org/unsd/publication/seriesm/seriesm_88e.pdf

* <http://dma.jrc.it/services/fuzzyg/>

** https://geonames.usgs.gov/apex/www_flow.accept

Класификации и категоризации

Предметни рубрики (Subject Headings). Предметните рубрики са класификационна схема, която представлява набор от контролирани термини за отразяване темите на отделните документи в колекция. Обикновено в предметизационна схема се отразяват отношения на еквивалентност (синонимия) и йерархия между термините в списъка (речника). Речникът може да бъде представен в азбучен ред или чрез ограничена йерархична структура^{***}. При използване отделните предметни рубрики по правило се свързват (координират), според специални правила, за да изразят понятия, които са по-конкретни (тесни или специфични). Примери: *Medical Subject Headings (MeSH)*, *Library of Congress Subject Headings (LCSH)*.

Класификационни и категоризационни схеми, таксономии, фолксономии (Classification Scheme, Categorization Scheme, Taxonomie, Folksonomy). Тези термини често се използват като синоними, макар че между отделните системи, които обозначават, има разлики¹²¹. Общото между тях е, че се основават на подхода разпределяне на обектите по един или няколко общи признаци в класове или категории. Класификационните системи за организация на знания осигуряват йерархично подреждане с цифрова и/или азбучна нотация, за да представят широки теми и по тесни понятия. В тяхната структура е заложена йерархичната връзка между понятията. Примери за класификационни схеми: *Library of Congress Classification Schedule*, *Dewey Decimal Classification*, *Универсална десетична класификация*.

Таксономията е набор от понятия, структурирани йерархично на основата на класификационна схема. Йерархични релации са само от вида „род-вид” и „част-цяло”.

Категоризационните схеми или тематичните категории често се използват за групиране на термини от тезаурус в по-широки теми, напр. тематични категории на сайтове¹²².

Фолксономията е социална класификация, основана на съвместно създаване и управление на етикети (тагове) от потребителите за поясняване и категоризиране на съдържание. Тагването, което е една от определящите характеристики на Web 2.0 услугите, позволява на потребителите заедно да класифицират и намират информация. Целта на тази класификация улесняване на намирането на информация по разбираем за потребителите

^{***} Често предметизационни схеми са развити в тезауруси (отразяват различни семантични отношения еквивалентни, йерархични, асоциативни), за да служат в електронна среда.

¹²¹ Hodge, Gail. Op. cit., p. 6.

¹²² Ibid.

начин. Визуално фолксонимията се изобразява като „облак” от етикети в уебсайта като таговете използвани по-често се изобразяват с по-едър шрифт.

Системи от термини/понятия и семантични връзки:

– Тезауруси (Thesauruses). Тезаурусите се основават на понятия (концепти) и показват взаимоотношения между термините, които ги обозначават. Взаимоотношенията, които обикновено се изразяват в тезаурус са йерархия, еквивалентност (синонимия) и асоциация или свързаност. Тези взаимоотношения обикновено се представят от нотация BT (по-широк термин), NT (по-тесен термин), SY (синоним) и RT (асоциативен или свързан термин). Връзките между термините могат да бъдат в различна степен развити в конкретни системи, например, в Unified Medical Language System (UMLS) на Националната медицинска библиотека (National Library of Medicine) са определени повече от 40 взаимовръзки, повечето от които са асоциативни.

Международен стандарт регламентира стандартни взаимовръзки между термините, идентифицирането на предпочитани термини и правила за създаване на връзки*. Днес са разработени най-различни тезауруси по обем, приложение и тематичен обхват. Повечето са създадени за отделна научна област, дисциплина или конкретен продукт, напр.: *Aquatic Sciences and Fisheries Thesaurus* (на Организацията по храните и земеделието *Food and Agricultural Organization*), *Thesaurus for aeronautics and aerospace-related topics* (на Националното управление по въздухоплаване и изследване на космическото пространство, НАСА, САЩ), *The Art and Architecture Thesaurus* (на Музей на изкуствата на тръстовия фонд на Жан Пол Гети, САЩ).

– Концептуалните карти (concept maps) по същество са вид диаграми, показващи връзките между понятия в избрана област. Те са средство за графично представяне на знания, (напр. концептуалните карти в DELOS DLRM), но могат да бъдат разработени на език на семантичния уеб, напр. RDF, и тогава имат някои от характеристиките на неформалните онтологии. Те също могат да бъдат разглеждани като първа стъпка от разработването на онтологии, когато се решава кои понятия да бъдат включени и се определят връзките между тях.

– Тематичните карти (Topic maps) са инструмент за представяне на информация, предназначено най-често за улесняване на търсенето. Освен понятия и връзки между тях тематичните карти съдържат и препратки към ресурси. Тематичните карти могат да се

* ISO 25964-1:2011 Information and documentation -- Thesauri and interoperability with other vocabularies -- Part 1: Thesauri for information retrieval.

използват за изграждане на логически структури за интегрирането на знания, за създаване на системи, базирани на знания и др.

– Семантични мрежи (Semantic Networks). Възможността за автоматично обработване на документи (ресурси) на естествен език (natural language processing) става възможна с развитието на семантичните мрежи. Този вид система за организация на знанията структурира понятия и термини не в йерархия, а като мрежа (графи) от възли и дъги, които ги свързват. Връзките в семантичните мрежи обикновено надхвърлят стандартните за тезаурус синоним, по-широк, по-тесен термин или свързан термин. Те могат да бъдат и специфични като част-цяло, причина-следствие и др. Най-известната семантична мрежа е *WordNet*, която се използва в различни търсещи машини.

– Онтологии (Ontologies). Онтологиите са най-съвременната система за организация на знанието. Те се разработват от специалисти от областта на „управление на знанията“ като специфични концептуални модели. Те позволяват да се представят сложни взаимоотношения между обектите в предметната област и да се включат правилата и аксиомите за тяхното свързване, които липсват в семантичните мрежи. Онтологиите, които описват знания в определена област, често са свързани със системи за извличане на данни (data mining) и управление на знания (knowledge management)¹²³.

Всички тези примери на системи за организация на знанието, които се различават по сложност, структура и предназначение, могат да осигурят организация и подобрен достъп до ресурсите на електронната библиотека.

2.2. Съвременни уеб технологии и развитието на електронните библиотеки

Големият брой потребители на уеб със завишаващи се потребности и изисквания, както и количеството на уеб ресурсите, налагат появата и развитието на множество технологии за сравнително кратко време. За по-малко от 30 г. уеб преминава през три етапа на развитие (често наричани версии) – уеб 1.0, уеб 2.0, уеб 3.0¹²⁴.

Уеб 1.0 (Традиционен уеб) е термин, описващ състоянието на уеб през първите десет години от неговото съществуване – 90-те години на XX в. Характерни за този етап на развитие на уеб мрежата са статичните уебстраници. Съдържанието се разработва от специалисти по уеб дизайн, което е скъпо и не може да гарантира своевременна актуализация. Ролята на крайните потребители е пасивна. Предполага четене (използване)

¹²³ Hodge, Gail. Op. cit., p. 7.

¹²⁴ Подробно за информационната инфраструктура и развитието на уеб вж: Харизанова, Оля. *Библиотеките и мрежовото общество*. София: Унив. изд. Св. Климент Охридски, 2010, с. 132–169.

на съдържание без възможност за създаване на ново съдържание, изразяване на мнение. Форумите и чатове са единствената възможност за някаква форма на интерактивност. Търсенето е изцяло основано на синтактични принципи, а за обработка на информацията се използват статистически алгоритми. Най-ранната версия на уеб свързва информацията чрез технологията на хипервръзките.

Уеб 2.0 (Социален уеб) развива технологиите, ориентирани към подпомагане и стимулиране на активното участие на потребителите в съвместно генериране и използване на уеб съдържание¹²⁵, т.е. тази версия на уеб по-скоро свързва хора. Този процес започва от началото на новото хилядолетие постепенно с разработката на все по-лесни за използване графични интерфейси и услуги, позволяващи на потребителите да публикуват съдържание. Синтактичното търсене се допълва с т. нар. социално търсене, което се базира на опита и предпочитанията на множество потребители в предходни периоди от време, използвайки определени зависимости в поведението им¹²⁶. Осигуряването на възможност за поставяне на етикети (социално тагиране) на ресурсите от страна на потребителя провокира доброволното аотиране на уеб съдържание. На този етап на развитие уеб се социализира чрез възможността за групиране на свързани потребители в зависимост от техните интереси, съвместно ползване на съдържание и услуги. Основната характеристика на тази версия е именно масовото създаване и използване на уеб съдържание и услуги. В резултат бе натрупан голям обем от информация, която изисква нови средства за обработка, организиране и търсене. Това е главната посока, в която се налага да се развива следващата версия на уеб, изискваща създаването и използването на свършено нови решения¹²⁷. От самото създаване на мрежата нейният комуникационен модел не е бил променян съществено. Уеб ресурсите се съхраняват в мрежата, без да се извършва цялостна машинна обработка на съдържанието, нито на текстовите, нито на мултимедийните документи. Огромният непрекъснато и неконтролируемо нарастващ брой документи на естествен език създава все по-сериозни проблеми, свързани с намирането, извличането, представянето, интерпретирането и обработката на информацията. Синтактичните принципи за търсене на информация основани на малък брой ключови думи не са в състояние да осигурят необходимата пълнота и точност в процеса на търсене поради характерните за естествените

¹²⁵ **Харизанова**, Оля. Новите измерения на World Wide Web и предизвикателствата пред библиотеките [online]. В: *Годишник на Софийския университет „Св. Климент Охридски“*. Философски факултет. Книга библиотечно-информационни науки [онлайн]. 2, 2010, с. 123–162. [Прегледан 20.05.2018]. Достъпно от: https://research.uni-sofia.bg/bitstream/10506/316/1/OHarizanova_godishnik.pdf

¹²⁶ **Добрева**, Стефка. Социално търсене в библиотечни бази от данни чрез използване на системата REDIS. В: *Научни трудове на Русенския университет* [online]. 2014, 53, серия 3.2, с. 138–143 [Прегледан 20.05.2018]. Достъпно от: <http://conf.uni-ruse.bg/bg/docs/cp14/3.2/3.2-24.pdf>

¹²⁷ *Еволюцията на уеб платформите е разгледана в: Харизанова*, Оля. Цит. съч.

езици многозначност, синонимия, омонимия, контекстна зависимост и допълнителни езикови проблеми, възникващи в условията на разноезична среда. Основна причина за недостатъците на Уеб 1.0 и Уеб 2.0 по отношение на информацията е невъзможността машините да обработват по семантичен начин (с разбиране на смисъла) съдържание (текст, изображения, аудио, мултимедия). Информацията, която се дава за съдържанието чрез изолирани от своя контекст думи, е недостатъчна и неефективна при търсене, когато се налага да се направи логически извод или когато трябва да се интегрира информация от различни източници. Решението е към ресурсите в уеб да се прикрепят данни (метаданни), които дават цялостна характеристика на ресурсите и които могат да се обработват от машините. Тези метаданни трябва да представят и идентифицират ресурса еднозначно с цел търсене и откриване. Описанието на ресурсите чрез метаданни трябва да бъде чрез стандартни схеми и да бъде достатъчно гъвкаво, за да може информацията да се комбинира и свързва. Така е възможно да се конструират логически изводи от съществуващи данни и информация, които ще направят машините „мислещи”. Съществуват два основни подхода за съхранение на метаданните, необходими за работата на семантичния уеб:

- в самите документи новите версии на маркиращите езици разработени от W3C позволяват комбиниране на презентационно съдържание и метаданни;
- в специални уеб достъпни бази знания, предоставящи стандартизиран интерфейс, за да могат външни приложения да ги използват.

Обикновено приложенията в уеб използват метаданните, съхранявани чрез двата подхода, като чрез първия се осигурява важната информация за уеб обектите, а чрез втория се гарантира правилната ѝ интерпретация, за което се използват онтологии.

Уеб 3.0 (Семантичен уеб). Новата еволюция на уеб трябва да предложи методи за автоматично обработване на големи обеми информация и решение на проблеми, свързани с неяснотата и противоречивостта на информацията. Уеб 3.0, наричана още „семантична мрежа” е съвременният етап на развитие на уеб, който се характеризира основно с: семантични технологии; средства от областта на изкуствен интелект; формати за представяне на данни, които гарантират тяхната използваемост от различни приложения; разработване на технологии за комуникация между приложенията в интернет; осигуряване на лесен достъп до информацията на основа свързаните данни (Linked Data) и др.

Семантичният уеб е набор от стандарти и добри практики за споделяне на данни и семантика на данните по интернет с цел използването им от компютърните приложения в

процеса на обработка на информацията¹²⁸. Докато уеб 1.0 предполага съдържанието да се разработва от професионалисти, а уеб 2.0 разчита на общностите от равноправни потребители, то уеб 3.0 разчита на машините за обработката на информацията и извличането на знания, за да предложи на потребителите по-ефективно търсене на информация и знания, основано на логически извод, разсъждения и качество на информацията, която да оправдава тяхното доверие. Уеб 3.0 предлага семантично търсене, свързване и интеграция на данни, тяхното многократно използване в различни контексти на основа интелигентни и семантични технологии, базирани на логически формализми.

Според създателя на уеб Тим Бърнърс-Лий семантичният интернет представлява разширение на сегашния интернет, в което на информацията е даден добре дефиниран смисъл, позволяващ на хора и компютри да работят съвместно по-добре от преди¹²⁹.

Семантичният уеб е концепция за развитието на уеб, която има за цел да представи информацията в мрежата във вид, улесняващ компютърната ѝ обработка чрез използване на методи и средства, базирани на нейния смисъл, а също и да предложи подходящ софтуер за това. В този аспект семантичният уеб представлява единна система от архитектура, технологии, стандарти и инструментални средства за представяне, обработка, структуриране, търсене, извличане и достъп до информацията по нов начин. Основната цел на технологиите на уеб 3.0 е осигуряването на методи и средства за автоматизация на обработката на информацията с цел лесното ѝ намиране и използване. В случая информация се използва в широк смисъл – статични уеб страници, уеб документи, информация от базите от данни, информация за потребителите (явна информация, съхранена в потребителски профил, и информация, произтичаща от действията на потребителя в уеб) и т.н.

2.2.1. Технологиите на Семантичния уеб

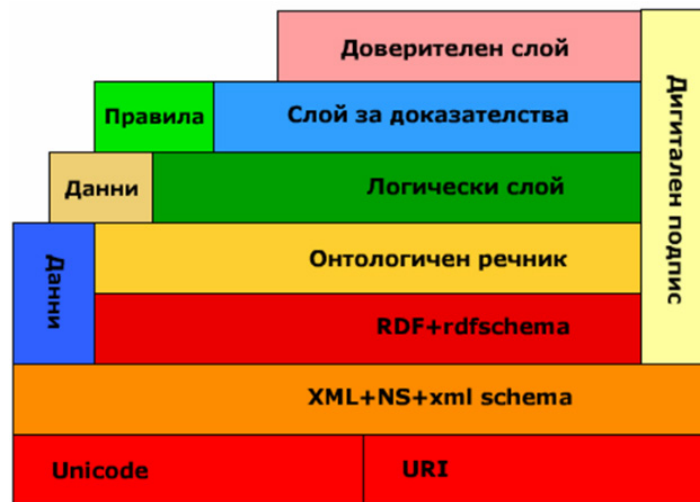
През 2001 г. създателят на уеб представя т. нар. „технологичен стек“ на семантичния уеб (semantic web stack)¹³⁰. Той се реализира чрез няколко взаимосвързани уеб стандарта, специално разработени за постигане на семантична обработка на информацията (фиг. 13).

¹²⁸ **Иванова**, Татяна. Глобалната семантична мрежа..., с. 21.

¹²⁹ **Berners-Lee**, Tim, **Hendler**, James, **Lassila**, Ora. The Semantic Web: A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. In: Scientific American [online]. May 2001 [Viewed 28.06.2018]. Available from: https://www-sop.inria.fr/acacia/cours/essi2006/Scientific%20American_%20Feature%20Article_%20The%20Semantic%20Web_%20May%202001.pdf; **Berners-Lee**, Tim. Semantic Web on XML [online]. Washington DC 2000/12/06. Slides [Viewed 28.06.2018]. Available from: <http://www.w3.org/2000/Talks/1206-xml2k-tbl>

¹³⁰ *Архитектурата на семантичния уеб и отделните компоненти са представени в: Цокева, Весела. Семантичната мрежа – от намиране на информация към получаване на знания. В: През призмата на информационната наука. София: Унив. изд. „Св. Климент Охридски”, 2016, с. 175–192.*

Фигура 13. Визия на Тим Бърнарс-Лий за семантичния уеб.



Най-долните два слоя съдържат технологии, които са добре известни от хипертекстовата мрежа (Уеб 2.0) и без промяна дават основа за семантичната мрежа:

Unicode (Уникод) е стандарт за кодиране, представяне и обработка на текст на повечето писмености в света. Съдържа набор от символи, използвани от различните азбуки в света, включително специални символи в науката. Стандартът се използва в множество съвременни технологии, включително XML.

URI (Uniform Resource Identifier) е уеб стандарт, определящ формата за идентифициране на ресурсите* и имената в семантичния уеб, (както и в целия интернет), използван за достъп до всякакви уеб ресурси. Има два вида идентификатори на ресурси: имена на ресурси (Uniform Resource Name, URN) и адреси на ресурси (Uniform Resource Locator, URL). С развитието на интернет вече се използват интернационализирани идентификатори на ресурси, които използват Unicode кодиране и имена на идентификатори на различни естествен език.

XML (eXtensible Markup Language) е стандарт (метаезик) за съхраняване, обмен и предаване на данни. Той е в основата на семантичния стек и се използва като синтаксис и базов език за разработване на други езици и технологии в уеб, напр. описание на онтологии, уеб услуги и аотиране на уеб съдържание¹³¹. Други технологии свързани с XML са:

* Тук ресурс се определя като всичко, което има идентичност, съгласно интернет стандарт RFC 2396 Унифицирани идентификатори за ресурса: Общ синтаксис” (URI – „Uniform Resource Identifiers: Generic Syntax) <https://www.ietf.org/rfc/rfc2396.txt>

¹³¹ Повече за основните технологии на семантичния уеб XML и RDF виж в: Миланова, Милена. Библиографски данни в семантичния уеб. В: Годишник на Софийския университет „Св. Климент Охридски”. Философски факултет. Книга библиотечно-информационни науки [online]. Т. 7, 2015, с. 201–214. [Прегледан 23.06.2018]. Достъпно от: https://research.uni-sofia.bg/xmlui/bitstream/handle/123456789/1381/7_Milena%20Milanova.pdf?sequence=1

- *XML Schema* определя на допустимите тагове, структурата и съдържанието на конкретен XML документ;
- *XML Namespace* [<https://www.w3.org/TR/xml-names/>] – веб стандарт за определяне на имена на типове елементи и на атрибути, използвани в XML документи за еднозначното им идентифициране чрез URI;
- *XML Query* [<https://www.w3.org/XML/Query/>] – за извличане и комбиниране на информация от XML документи от бази от данни, от уебстраници и др.

Средните три слоя съдържат технологии, стандартизирани от W3C, за да позволят изграждането на семантични уеб приложения:

*RDF*** (Resource Description Framework) е веб стандарт за описание и обмен на данни между уеб приложения с машинноинтерпретируем синтаксис. На него се базира представянето на знанията във вид удобен за компютърна обработка, тъй като дава формална семантика като представя данните като наредени тройки „субект-предикат-обект“. RDF служи за описание на твърдения за уеб ресурси. Представянето на твърдения става чрез разклонен граф, където върховете са субекти или обекти, а дъгите са свойства на първите (предикати). Субектите и обектите в RDF задължително се означават с идентификатори URI, а предикатите могат да са и литерали.

*RDF Schema** разширение на RDF стандарта, което осигурява възможности за йерархична организация на уеб ресурсите и може да бъде използван за разработка на прости онтологии. XML и RDF схемите осигуряват преобразуването от и към почти всички използвани в момента стандарти за данни. Като техен недостатък се сочи минималната им възможност за изразяване на логически отношения.

Онтология (Ontology). Една от основните технологии за съхраняване и обработка на знания в Семантичния уеб са онтологиите. Те са ефективна информационна технология за изразяване на семантични отношения между понятия във формат, удобен преди всичко за компютърна обработка, но директно четим и използваем и от хората. На основата на RDF и RDF Schema е разработен специален език за разработване на онтологии, позволяващ изразяване на по-сложни семантични релации между обектите, базиран на дескриптивната логика *OWL* - Web Ontology Language**.

** <https://www.w3.org/RDF/>

* <https://www.w3.org/TR/rdf-schema/>

** <https://www.w3.org/OWL>

Логика (Logic). Езиците за описание на онтологии се обогатяват с допълнителни логически правила с цел описание на по-сложни зависимости между данните, лесно и ясно дефиниране на ограничения, описание на действия, които трябва да се извършат с данни от базата при настъпване на предварително зададени условия. Примери за езици, базирани на правила: *RuleML*^{***} (Rule markup language) – XML базиран обектно-ориентиран език за логическо програмиране с правила, предназначени за семантичния интернет и *SWRL*^{****} (Semantic Web Rule Language) – предложен като веб стандарт за обработка на знания, базиран на правила.

Последните два слоя са свързани с технологии, които все още не са стандартизирани, или са само идеи, които трябва да бъдат изпълнени, за да се реализира семантичния веб.

Доказателства (Proof). Нивото на доказателствата включва самия дедуктивен процес за извеждане на нови знания, а също така и проверка на легитимността на източника на информация на основа цифров подпис.

Доверие (Trust). Доверието има два аспекта: а) дали информацията идва от доверен източник и б) дали полученото ново знание е основано на формална логика. Доверието обаче все още се свързва с известен субективизъм, тъй като може да се основа на предположения относно бъдещото поведение на потребителите, основано на техните минали действия. За да се осигури доверие между потребителите, трябва да има стандарт за знанията в семантичния веб, на чиято база да се извършва филтриране на информацията.

Този ред на семантични технологии не е изчерпателен, като се има предвид процеса на развитие на Семантичния веб. Предизвикателствата са фокусирани върху слоевете – доказателства, като целта е да се постигне доверие. Потребителите трябва да разчитат на сигурност, коректност и качеството на събраната и извлечена информация, за да се постигне ниво на доверие¹³². Затова значение имат средствата за преобразуване на данни (криптиране), с оглед прикриване на тяхната семантика, предотвратяване на неоторизиран достъп или тяхната промяна (*Cryptography*).

Тъй като за семантичната обработка на представената чрез онтологии информация е необходим качествено нов софтуер, то той се дискутира от някои изследователи като

^{***} <http://ruleml.org/index.html>

^{****} <https://www.w3.org/Submission/SWRL/>

¹³² Цокева, Весела. Цит съч., с. 188.

компонент в контекста на цялостната архитектура¹³³. Основа за реализиране на концепцията за семантичния уеб е изграждането на онтологии^{134,135}.

2.2.2. Онтологиите като средство за представяне на знанията в уеб

Основната част от ресурсите в интернет в момента са в разноезиков текстов формат. В този си вид те са предназначени за директно използване от човека. До изграждането на семантичен уеб компютрите имат ограничено участие в обработката на информацията в интернет. Функциите им се свеждат основно до съхранение и търсене. Същевременно поради неимоверното нарастване на обема на информацията в интернет, търсенето, основано на синтактични принципи (по ключови думи от текста на предварително индексирани от търсещи машини документи) се определя като неефективно. Не е възможно да се постигне нито достатъчно ниво на пълнота, нито пък точност в процеса на търсене¹³⁶.

Решение се търси в използването на онтологии. Използването на онтологии позволява всеобщото разбиране на знанията в определена предметна област, компютърното му обработване и извеждането на изводи (нови знания). В съвременните информационни системи, в частност електронните библиотеки, потребителят все по-често се поставя в центъра и системата прави опити максимално да задоволи изискванията, потребностите и желанията му с цел да го привлече за дълго време¹³⁷. Това предполага осигуряване на механизми за бързо и лесно търсене, семантично описание на различни типове ресурси, извличане на ново знание, съпоставяне на метаданните за информационни обекти с нуждите на потребителя, правене на интелигентни изводи по отношение на най-подходящи ресурси за всеки отделен потребител и др.

По отношение на информационното съдържание онтологиите могат да решат проблема с различния начин на възприемане на семантиката на понятията чрез установяване на споделена концептуализация в предметна област. Онтологии се използват и при организиране на процеса на преглеждане на документите. Те задават структурни отношения между съставни части на един информационен обект (документ) или с други

¹³³ **Иванова**, Татяна. Глобалната семантична мрежа..., с. 26.

¹³⁴ **Gärdenfors**, Peter. *How to make the Semantic Web more semantic* [online]. [Viewed 20.06.2018]. Available from: <https://slab.org/tmp/Gardenfors04.pdf>

¹³⁵ **Gruber**, T. R. Collective Knowledge Systems: Where the Social Web meets the Semantic Web. In: *Journal of Web Semantics* [online]. 2008, V. 6, № 1, p. 4–13. [Viewed 28.06.2018]. Available from: <http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=D4138B8FE67DEB6EEA61CCF8FC9E70B8?doi=10.1.1.106.183&rep=rep1&type=pdf>

¹³⁶ **Иванова**, Татяна Иванова. *Анализ и разработка на семантични технологии за търсене в уеб: Дисертация*. София: ТУ-София, 2009, с. 10.

¹³⁷ Пак там, с. 13.

обекти чрез релации като Next, Prev, IsPartOf, HasPart, References, IsReferencedBy, IsBasedOn, IsBasisFor, Requires, IsRequiredBy¹³⁸.

Онтологии се използват и за да опишат потребителите, което предоставя възможност за т.нар. адаптивно търсене (с отчитане на индивидуалните специфики и личния профил на потребителя).

Най-общо онтологията се прилага за подобряване на информационното търсене (разширяване на заявките чрез онтология на предметната област и онтология на потребителите); задаване на общ смисъл на метаданните за информационните ресурси (библиографски онтологии); прилагане на обща база знания от различни системи, основани на знания (базови онтологии)¹³⁹.

В последните години онтологията придобиха широко практическо приложение – от комуникациите между хора и компютърни системи до автоматично извеждане и многократно използване на знания¹⁴⁰.

В научната литература има много дефиниции на понятието онтология, които поставят акцент върху различни аспекти: абстрактно-теоретичен (насочен към формалното определение на понятието); практически, свързан с определяне на предназначението на онтологията; структурирен, представящ компонентите на онтологията и др.

Най-общо онтологията дава формални наименования и определя типовете обекти (класовете), обектите (екземплярите) и техните свойства (атрибути) и взаимовръзките между тях. Най-често цитираният изследовател Т. Грубер определя онтологията като „явна спецификация на концептуализация“¹⁴¹. Това определение е уточнено от В. Борст „онтологията е формална спецификация на съгласувана концептуализация“¹⁴². Тези свойства на онтологията се определят така:

¹³⁸ Павлова, Лилия Радославова. *Семантични технологии и средства в цифрови библиотеки за художествено културно наследство: Автореферат* [online]. София, 2013, с. 13 [Прегледан 20.06.2018]. Достъпно от: http://www.math.bas.bg/IMIdocs/ZRASRB/docs/Lilia_Pavlova/LPavlova-avtoreferat.pdf

¹³⁹ Kruk, Sebastian Ryszard, Haslhofer, Bernhard, Piotrowski, Piotr, Westerski, Adam, Woroniecki, Tomasz. *The Role of Ontologies in Semantic Digital Libraries* [online]. [Viewed 20.06.2018]. Available from: https://www.researchgate.net/publication/242734216_The_Role_of_Ontologies_in_Semantic_Digital_Libraries

¹⁴⁰ Uschold, Mike, Gruninger, Michael. *Ontologies Principles, Methods and Applications*. In: *Knowledge Engineering Review* [online]. Vol. 11, № 2, June 1996 [Viewed 20.06.2018]. Available from: <http://www.aii.ed.ac.uk/publications/documents/1996/96-ker-intro-ontologies.pdf>

¹⁴¹ Gruber, T. R. A Translation Approach to Portable Ontology Specifications [online]. In: *Knowledge Acquisition*, 1993, 2, 199–220. [Viewed 20.06.2018]. Available from: <http://tomgruber.org/writing/ontologia-kaj-1993.pdf>

¹⁴² Цит. по: Guarino, Nicola, Oberle, Daniel, Staab, Steffen. What Is an Ontology? In: *Handbook on Ontologies* [online]. Berlin-Heidelberg: Springer, 2009. [Viewed 20.06.2018]. Available from: https://www.researchgate.net/publication/226279556_What_Is_an_Ontology

Формална спецификация – точно, детайлно описание представено на формален език (напр. RDF), който в концепцията за семантичния уеб е машинночетим (и обработваем от програма).

Концептуализация – структура (абстрактен модел) на част от реалността, разглеждана независимо от речника на предметната област или конкретна ситуация. Тя се състои от множество от обекти (понятия) и връзки между тях.

Друго определение¹⁴³ описва структурата на онтологията, определена от нейните съставни компоненти – $O_{\text{онтология}} = (C, P, Hc, Rc, Hr, I, Ri, A)$, където:

C (от „concepts”) – множество от понятия в предметна област (класове от понятия);

P (от „properties”) – множество от свойства на основата, на които понятията могат да принадлежат към клас;

Hc (от „hierarchy concepts”) – йерархия, в която са структурирани понятията;

Rc (от „relationships concepts”) – множество от връзки между понятията;

Hr (от „hierarchy relationships”) – йерархия на връзките между понятията;

I („individuals”) – отделни екземпляри;

Ri (от „relationships individuals”) – връзки между отделните екземпляри;

A („axioms”) – аксиоми (правила) за извеждане на нови знания от обработваните.

Поради комплексността на онтологиите и възможността те да бъдат разглеждани в няколко различни аспекти (на знанията, които представят; на логиката, която е в основата им; или на информационните технологии, с които се обработват) в литературата се използват термини от различни терминосистеми. Когато контекстът е представяне на онтологията на език за програмиране се използват термините типове, обекти и отношения. Разглеждани като знания от предметна област, се наричат понятия, екземпляри, свойства, правила. На езика на дескриптивната логика съответстват класове, инстанции, роли, аксиоми.

Дефиницията на структурата на онтологиите показва, че те са сложни системи от няколко типа компоненти. Компонентите на онтологията могат да се разделят на базови и допълнителни.

а) Базови компоненти на онтология:

Клас – обобщение (генерализация) на характеристиките, които притежават набор от екземпляри. Един клас може да бъде произволно абстрактно понятие, подходящо за целта на конкретна онтология, напр. процедура, функционалност, действие, стратегия и т.н.

¹⁴³ **Иванова**, Гатяна. Глобалната семантична мрежа..., с. 32.

Например, термините „Софийски университет” и „Станфордски университет” не са класове, а екземпляри, докато терминът „Университет” е абстракция, която събира в клас различни екземпляри. Класовете притежават име, описание на естествен език и множество релации (слотове), които ги характеризират. Един клас се свързва с едно или повече синоними. Свързва се с другите класове чрез релации “йерархия” или “асоциация”.

Релации представят взаимните връзки между два и повече класа в онтология. Релациите са йерархични и нейерархични. Йерархичните релации между класовете определят таксономията (род-вид) и мерееологията (цяло-част) в една онтология. Нейерархичните релации определят останалите връзки между елементите на онтологията и се наричат още асоциативни релации.

Греъм Хърст определя следните основни релации в онтологиите:

– синонимност (еквивалентност): релации между термини, които имат едно и също значение, а един от тях се означава като препоръчителен термин (preferred term, PT); останалите са синоними, наричани още неепоръчителни термини (non-preferred terms, NPT). Идентичните термини образуват синонимно множество или (synsets, synonym sets, synonym rings);

– таксономични релации: хипонимия (вид или подклас) и инверсната ѝ хиперонимия (род или суперклас), изграждат семантичната йерархия, (наричана още йерархия Is A/Has A). Таксономичните релации описват йерархичната структура на класовете, т.е. разделянето на главните класове (суперкласове) на подчинени класове (подкласове). Релациите (слотовете) на суперкласовете в тази структура се наследяват от подкласовете.

– мерееологични релации: меронимия (цяло) и инверсната ѝ холонимия (част) също изграждат йерархии от тип част-цяло. Те изразяват ситуации в онтологиите, при които цялото е съставено от части, разглеждани като отделни класове. Частите участват в цялото в различни отношения, напр.¹⁴⁴: компонент/обект (обектът е съставен от компоненти, всеки от които има собствена функционалност – напр. колелата са части от автомобила); член/колекция (членовете не играят функционална роля по отношение на цялото и могат да бъдат отделени от него, например дървото е част от гората) и др.

¹⁴⁴Artale, A., Franconi, E., Guarino, N., Pazzi, L. Part-Whole Relations in Object-Centered Systems: An Overview. In: *Data and Knowledge Engineering*[online]. Vol. 20, 1996, 3, p. 347–383. [Viewed 20.06.2018]. Available from: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.33.131&rep=rep1&type=pdf>

– полийерархия – релации на клас, който има множество подкласове и суперкласове¹⁴⁵.

Релациите винаги са двойки – всеки тип релация може да се разглежда като двойка от правата и инверсната си форма (напр. Is A/Has A, Is Part of/Has Part и др.).

Екземпляри – отделни инстанции, които илюстрират класовете. Те са конкретни обекти, притежаващи характеристиките на класа, с който са свързани. Онтологията може да представя само абстрактни знания в дадена област, изразени чрез понятия (класове), връзки между тях и аксиоми, но може да съдържа и конкретни екземпляри.

б) Допълнителни елементи на онтология:

Аксиоми (правила) – декларативно представено знание, което се приема за вярно без проверка¹⁴⁶. Те изразяват общоприети твърдения, отнасящи се до класовете, релациите и екземплярите. Тези твърдения се използват за логическо извеждане на други твърдения (нови знания). Аксиомите са два вида: структурни и неструктурни. Структурните аксиоми ограничават структурата на онтологията, напр. „клас А и клас В са непресичащи се”, означава, че един екземпляр не може да принадлежи едновременно на двата класа. Неструктурните аксиоми са локални за един клас и ограничават неговата интерпретация, изразявайки състояния на неговите атрибути, например, стойностите на атрибутите сила, маса и ускорение за всеки екземпляр от клас „твърдо тяло” се подчиняват на зависимостта „сила = маса X ускорение”¹⁴⁷.

Функции – функциите са по-сложни релации (напр. Consist of или Is A List of са функции, защото за разлика от бинарните релации, свързват множество елементи класове, екземпляри и атрибути).

2.2.2.1. Типове онтологии

Онтологиите могат да бъдат класифицирани в зависимост от различни признаци: областта, за която се отнасят; предназначение; език; нивото на детайлност на представените знания; предмета на концептуализация; използваните за обработката формализми;

¹⁴⁵ Hirst, Graeme. *Ontology and the Lexicon* [online]. [Viewed 20.06.2018]. Available from: <ftp://ftp.cs.toronto.edu/pub/gh/Hirst-Ontol-2009.pdf>

¹⁴⁶ Mizoguchi, R., Sinitsa, K., Ikeda, M. *Knowledge Engineering of Educational Systems for Authoring System Design: A preliminary results of task ontology design* [online]. [Viewed 20.06.2018]. Available from: <http://www.ei.sanken.osaka-u.ac.jp/pub/miz/miz-eaied96.html>

¹⁴⁷ Делийска, Боряна. *Електронно обучение, базирано на онтологии, в областта на компютърните мрежи и екоинформатиката: Хабилитационен труд* [online]. София, 2008. [Прегледан 20.06.2018]. Достъпно от: https://www.researchgate.net/publication/319507476_Elektronno_obucenie_bazirano_na_ontologii_v_oblastta_na_komputrните_mrezi_i_ekoinformatikata

логическата сложност и др.^{148,149}. Често дискутирани в специализираната литература са следните признаци и категории онтологии:

а) Според предмета на концептуализация се формулират два основни типа онтологии - общи и предметни. Други изследователи предлагат по-подробна класификация¹⁵⁰:

Общите онтологии (common, general, core, generic ontologies, meta-ontologies) са минимални онтологии, съдържащи общи, използвани в много предметни области, понятия като пространство, време, процес и т.н., които са независими от конкретна задача или предметна област. Те съдържат унифицирани, представени явно и използвани в големи общности от потребители понятия с абстрактни дефиниции. Някои онтологии от този тип са: Basic Formal Ontology, CYC, General Formal Ontology и др., които могат да се използват в различни приложения.

Предметните онтологии (domain ontologies) представляват формални модели на определена област от знания (в някоя наука, общност, приложение и т.н.) и съдържат понятията от тези области. Предметните онтологии трябва да бъдат многократно използвани в конкретната област. В много дисциплини се разработват онтологии - концептуализация на основните знания, които могат да се използват от експертите в областта за различни цели. Например в библиотечно-библиографската област такива онтологии са: FRBR, IFLA LRM, CIDOC и др.

Онтологията за представяне на знания (представителни онтологии) (knowledge representation ontologies, representation ontologies) служат за създаване на рамка за представяне на знания въобще, без претенции за конкретно приложение.. Пример за представителна онтология е Frame Ontology, която се използва в езика Ontolingua¹⁵¹. Тя дефинира термини като релация, функция, клас и др., използвани за моделиране на онтологии.

Приложните онтологии (application ontologies) обхващат необходимото знание за моделиране на конкретно приложение. Те се разработват за решаване на конкретна приложна задача. Те представят знанията в контекста на спецификата на приложението и само частта от предметната терминология, важна за него¹⁵². Обикновено приложните онтологии са смес от класове, взети от предметни и общи онтологии. Примери за приложни

¹⁴⁸ **Иванова**, Татяна. Глобалната семантична мрежа..., с. 9.

¹⁴⁹ **Делийска**, Боряна. Цит. съч.

¹⁵⁰ **van Heijst**, G., **Schreiber**, G. *Wielinga: Using Explicit Ontologies in Knowledge-Based System Development* *Explicit Ontologies in Knowledge-Based System Development* [online]. [Viewed 20.06.2018]. Available from: https://www.researchgate.net/publication/243779867_Wielinga_Using_Explicit_Ontologies_in_Knowledge-Based_System_Development

¹⁵¹ **Gruber**, T.R. A Translation Approach to Portable ...

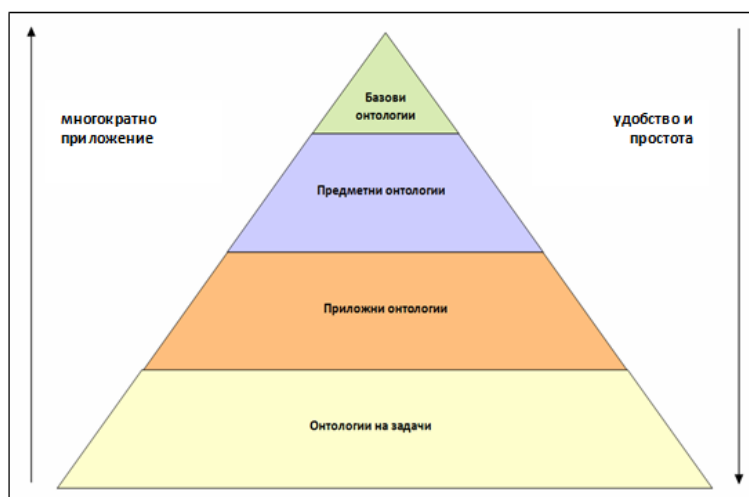
¹⁵² **Иванова**, Татяна. Глобалната семантична мрежа..., с. 119.

онтологии са онтологии на информационни системи в дадена предметна област и т.н. Поради тясната си обвързаност с приложението, за което са разработени, те са твърде специфични и не могат да се използват многократно, без преработване.

Онтологията на задачи и методи (task ontologies, method ontologies) предлагат систематични речници на термини с дефиниции на класове и релации между тях, свързани с решаването на специфични задачи (зависими или независими от предметната област). Такива са онтологии за проектиране на сграда, парк, лаборатория или онтология за събиране на данни и т.н.¹⁵³

Специално внимание¹⁵⁴ се обръща на зависимостите между конкретизацията на онтологията и възможността за повторно/многократно приложение на онтологията, от една страна, и удобството и простота на работа с нея, от друга, отразени на фиг 14. Колкото по-конкретна е една онтология (на задачи, на приложение, на предметна област), толкова по-малко са възможностите за многократно използване в други приложения.

Фигура 14. Зависимост между приложение и удобство на работа на типове онтологии



б) Според степента на аксиоматизация:

Съгласно аксиоматизацията на дефинициите в онтологията¹⁵⁵ и според типа на включения в тях речник на понятията, онтологията са формални, полужформални, полунеформални и неформални (терминологични, лексикални).

¹⁵³ Делийска, Боряна. Цит. съч.

¹⁵⁴ Хой, Ле. Исследоване и разработка електронных библиотек на основе явного описания семантики ресурсов с использованием технологий Semantic Web: Диссертация. Томск, 2013, с. 34.

¹⁵⁵ Sowa, John. *Knowledge representation: Logical, philosophical, and computational foundations* [online]. [Viewed 28.06.2018]. Available from: https://www.researchgate.net/publication/225070439_Knowledge_Representation_Logical_Philosophical_and_Computational_Foundations

Формалните онтологии са с точно дефинирани термини и формално описана семантика с правила за пълнота и съответни доказателства. Чрез програмен език те автоматично се транслират в логика. В тях е възможно да се изпълняват сложни изводи, поддържани от логическите им основи, да се проверява съвместимостта по време на създаването им и да се произвеждат нови знания по време на заявки¹⁵⁶. Пример за формална онтология е Сус*.

Полуформалните онтологии са изразени с термините на изкуствено създаден език¹⁵⁷.

Полунеформалните онтологии се описват в ограничена и специално структурирана форма на естествения език.

Неформалните онтологии са изразени в свободна форма на естествен език и не се нуждаят от аксиоми, ограничаващи използването на класовете им. Пример за такива онтологии е WordNet.

Теоретично разликата между неформалните и формалните онтологии е в степента на аксиоматизация. Колкото повече аксиоми се добавят към терминологична онтология, толкова повече тя се превръща във формална¹⁵⁸.

в) Според използвания естествен език:

В зависимост от езика, използван в онтологията, на който е представена терминологията, онтологиите могат да бъдат несвързани конкретно с никой естествен език (формални онтологии, чиито елементи са означени с абстрактни термини или номера), едноезични (такива са почти всички онтологии, като основната част от тях са разработени на английски език), или многоезични, представящи термини на няколко езика (такива започнаха да се разработват през последните години и все още са само резултати от изследвания и проекти). Освен имената на елементите като термини на естествен език едноезичните или многоезичните онтологии могат да съдържат и друга лексикална информация като синоними, антоними, части на речта и др. Такива са лексикалните онтологии, създавани от компютърната лингвистика за машинен обработка на текст и превод.

¹⁵⁶ Делийска, Боряна. Цит. съч.

* <http://www.Cyc.com>

¹⁵⁷ Пак там.

¹⁵⁸ Делийска, Боряна. Цит. съч., с. 78

2.2.2.2. Приложения на онтологии в електронна библиотека

Съвременните електронни библиотеки представляват важна част от пространството на необхватния уеб, което съдържа организирана качествена информация. Управлението на тази информация и споделянето ѝ между различни информационни системи и/или електронни библиотеки все още е свързано с някои сериозни предизвикателства: ниска степен на семантична оперативна съвместимост и неефективно търсене¹⁵⁹. Решения на тези предизвикателства се търсят в постиженията на семантичния уеб и в частност в прилагането на онтологии¹⁶⁰.

Предизвикателство 1: Семантична съвместимост.

За съвременните електронни библиотеки възможността за взаимодействие е задължителна. Използвайки протоколи „Peer-to-peer” (P2P)*, могат да бъдат интегрирани отделни електронни библиотеки в една голяма библиотека. За да се постигне семантична съвместимост е необходимо споделено използване на онтологии (базова, предметни и др.). Онтологията може да бъде „общ знаменател”, инструмент, който решава проблемите, свързани с взаимодействието на библиотеките на семантично ниво.

Осигуряването на семантична съвместимост е най-важната, но и най-трудна задача. Тя има два основни взаимоотношения – първият е свързан с съгласуваността на елементите на метаданните, а вторият – с присвояването на значения на елементите на метаданните, което предполага използване на общоприети контролирани речници, тезауруси, нормативни файлове, онтологии. Съгласуваност на елементите на метаданните се постига чрез използването на стандартни схеми за описание и използването на библиографска онтология, която да осигурява превода между различните схеми. Присвояването на значения на елементите на метаданните става чрез използване на онтологии като средство за споделяне на общ речник в дадена област на знанието.

¹⁵⁹ **Synak**, Marcin, **Kruk**, Sebastian Ryszard. *MarcOnt Initiative the Ontology for the Librarian World* [online] [Viewed 24.06.2018]. Available from:

<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.114.3986&rep=rep1&type=pdf>

¹⁶⁰ **Kruk**, Sebastian Ryszard, **Synak**, Marcin, **Zimmermann**, Kerstin. *MarcOnt-integration ontology for bibliographic description formats* [online]. [Viewed 20.06.2018]. Available from:

https://www.researchgate.net/publication/228867665_MarcOnt-integration_ontology_for_bibliographic_description_formats

* *Peer-to-peer* – разпределена мрежова архитектура, съставена от участници, които отдават част от своите ресурси директно на разположение на другите мрежови участници без необходимост от централна мрежова инстанция (като сървъри или хостове). Участниците са едновременно „доставчици” и „потребители” на ресурси. *Peer-to-peer* технологиите позволяват интеграция на услуги за електронни библиотеки и споделяне на информация.

На база преглед и анализ на публикации, посветени на изследователския проект JeromeDL за използване на семантични технологии в електронни библиотеки¹⁶¹, идентифицирахме три области на приложение на онтологии:

– „надграждане” на библиографските метаданни на семантично ниво до машинноинтерпретируемо (чрез задаване на семантика на понятията, дефинирани в онтология);

– прилагане на онтология за моделиране на структурните аспекти на обектите, съхранявани в библиотечните системи;

– въвеждане на онтологиите като механизъм за описание на назначенията за потребителите и общностите, така че електронните библиотеки да се развият от системи със статична информация до динамични пространства на знанието¹⁶².

За постигане на семантична съвместимост между електронни библиотеки се използват онтологии с различни функции:

а) *Библиографски онтологии*. Обикновено всяка електронна библиотека използва определен формат (стандарт, схема) за метаданни с цел стандартно организиране на библиографската информация. Много популярен формат за описание на информационни ресурси и за обмен на библиографска информация, който произхожда от традиционните библиотечни системи за каталогизиране е MARC¹⁶³. Други широко използвани стандарти са Dublin Core, BibTeX¹⁶⁴ и др.

Тези формати обаче не задават семантиката на елементите (полетата) на метаданните по машинноинтерпретируем начин. За това е необходима онтология, която дава добре

¹⁶¹ *Литературни източници за JeromeDL*: **Kruk**, Sebastian Ryszard, **Mocan**, Adrian, **Sapkota**, Brahmananda, **Zaremba**, Michal. Building Semantic Web Services Infrastructure for Digital Libraries [online]. [Viewed 28.06.2018]. Available from: <https://pdfs.semanticscholar.org/1a7b/2118cbc7ff520d26a9397bbb2692be68f056.pdf>; **Kruk**, Sebastian Ryszard, **Decker**, Stefan, **Zieborak**, Lech. JeromeDL – Adding Semantic Web Technologies to Digital Libraries. In: *Database and Expert Systems Applications: 16th International Conference* [online]. DEXA 2005, Copenhagen, Denmark, August 22–26, 2005, Proceedings, p. 716–725. [Viewed 20.06.2018]. Available from: https://aran.library.nuigalway.ie/bitstream/handle/10379/644/70ae42e7_1.pdf?sequence=1&isAllowed=y; **Kruk**, Sebastian Ryszard, **Woroniccki**, Tomasz, **Gzella**, Adam, **Dabrowski**, Maciej. *JeromeDL – a Semantic Digital Library* [online]. [Viewed 20.06.2018]. Available from: <https://pdfs.semanticscholar.org/1a7b/2118cbc7ff520d26a9397bbb2692be68f056.pdf>; **Kruk**, Sebastian Ryszard, **Decker**, Stefan, **Zieborak**, Lech. *JeromeDL-Reconnecting Digital Libraries and the Semantic* [online]. [Viewed 20.06.2018]. Available from: https://www.researchgate.net/publication/228994759_JeromeDL-Reconnecting_Digital_Libraries_and_the_Semantic_Web

¹⁶² **Kruk**, Sebastian Ryszard, **Haslhofer**, Bernhard, **Piotrowski**, Piotr, **Westerski**, Adam, **Woroniccki**, Tomasz. The Role of Ontologies in Semantic ...

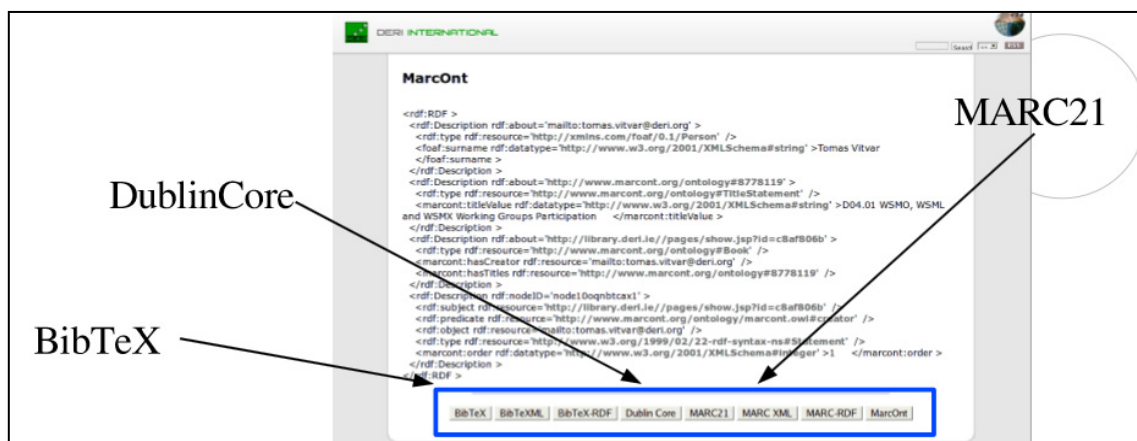
¹⁶³ **Миланова**, Милена. *Българската каталогизация в глобалното информационно пространство на XXI век. Анализи, стратегии, перспективи: Дисертация* [online]. София, 2008. [Прегледан 27.06.2018]. Достъпно от: https://research.uni-sofia.bg/bitstream/10506/1089/5/Milenana%20Milanova_diss.pdf

¹⁶⁴ **Synak**, Marcin, **Kruk**, Sebastian Ryszard. Op. cit.

дефинирани понятия и взаимовръзки между тях, т.е. семантика, и позволява да се извличат нови неявно изразени данни от библиографските описания.

Една от библиографските онтологии, която е разработена и се използва в семантичната електронна библиотека JeromeDL е MarcOnt Ontology¹⁶⁵. Тя има за цел да комбинира популярни стандарти за метаданни, които описват различни типове ресурси с различни нива на подробност – MARC 21, Dublin Core, BibTeX (фиг. 15).

Фигура 15. MarcOntOntology в JeromeDL



б) Структурни онтологии

Електронните библиотеки не само съхраняват и управляват библиографски метаданни и пълен текст, но трябва да представят структурно документите. В зависимост от типа на документа, съдържанието обикновено следва някаква структура, напр. типът документ „книга” се състои от отделни глави. Те могат да бъдат описани отделно като бъдат предоставени данни за свързаност с другите съставни части на книгата. Както показват изследвания¹⁶⁶ прилагането на онтологични описания на структурата на документите дава възможност за единен достъп до структурна и библиографска информация.

в) Онтологии на потребители

Обикновено библиотеката обслужва определени потребителски групи или общности. Индивидуалният потребител, обръщайки се към електронната библиотека може

¹⁶⁵ **Kruk**, Sebastian Ryszard, **Synak**, Marcin, **Zimmerman**, Kerstin. *MarcOnt – Integration Ontology for Bibliographic Description Formats* [online]. [Viewed 28.06.2018]. Available from: https://www.researchgate.net/publication/228867665_MarcOnt-integration_ontology_for_bibliographic_description_formats

¹⁶⁶ **Kruk**, Sebastian Ryszard, **Decker**, Stefan, **Zieborak**, Lech. JeromeDL – Adding Semantic Web Technologies to Digital Libraries. In: *Database and Expert Systems Applications: 16th International Conference* [online]. DEXA 2005, Copenhagen, Denmark, August 22–26, 2005, Proceedings, p.716–725. [Viewed 20.06.2018]. Available from: https://aran.library.nuigalway.ie/bitstream/handle/10379/644/70ae42e7_1.pdf?sequence=1&isAllowed=y

не само да търси и получава, но и да споделя информация и знания, като и да се учи от опита на други потребители. Например потребител може да се интересува какво чете и оценява като интересно лицето X, което потребителят цени като експерт в обща област на интереси. Възможността за взаимодействие се реализира чрез онтология, която описва потребителите. Такава онтология е представлява FOAF (Friend of a friend) Ontology*. Тя предоставя единен начин за описване на лица чрез свойствата: име, адрес на електронна поща и др. и позволява да се уточнят отношенията на свързаност („приятелство”)¹⁶⁷. Онтологията FOAF може да бъде разширена и да предлага споделяне на потребителски отметки (bookmarks) сред приятели, като по този начин прави възможно т.нар. филтриране чрез социалното семантично сътрудничество¹⁶⁸.

Представянето на потребителски профили на базата на FOAF позволява да се обвърже информацията за потребителите с други метаданни, така че потребителските профили и информацията за социалната мрежа да могат автоматично да се използват при търсене¹⁶⁹. Използването на общоприети онтологии и RDF като формат за данни позволява съхраняване и използване на информация за потребителите разбираеми за различни системи.

Предизвикателство 2: Подобряване на търсенето.

В резултат на несъвършената машинна обработка и синтактично търсене по ключови думи потребителят често получава от търсещата система незадоволителен резултат от информационното търсене (твърде голям или нулев резултат, най-подходящите резултати не са в началото на дългия списък, има само формални текстови съвпадения и др.).

Потребителите очакват съвременната електронна библиотека да е способна на персонализация и адаптивност особено в процеса на търсене и извличане на информация. Адаптивното търсене* е процес на селекция на най-подходящите резултати и тяхното подреждане в зависимост от конкретен потребител, група от потребители или всички потребители на системата. Адаптивното търсене предлага ред предимства пред търсенето по ключови думи като:

* <http://www.foaf-project.org/>

¹⁶⁷ **Kruk**, Sebastian Ryszard. FOAF-Realm – control your friends' access to the resource [online]. [Viewed 28.06.2018]. Available from: https://www.w3.org/2001/sw/Europe/events/foaf-galway/papers/fp/foaf_realm/

¹⁶⁸ **Kruk**, Sebastian Ryszard, **Decker**, Stefan. *Semantic Social Collaborative Filtering with FOAFRealm* [online] [Viewed 20.06.2018]. Available from:

https://aran.library.nuigalway.ie/bitstream/handle/10379/501/5e68d6a6_1.pdf?sequence=1&isAllowed=y

¹⁶⁹ **Kruk**, Sebastian Ryszard, **Decker**, Stefan, **Zieborak**, Lech. JeromeDL – Adding Semantic Web Technologies...

* *Адаптивното търсене се определя се още като семантично търсене, тъй като се основава на технологии и стандарти на семантичния уеб.*

- отчитане на индивидуалните специфики и личния профил на потребителя при оценяване и ранжирането на резултатите,
- отразяване на резултати от сходни търсения на други потребители,
- отчитане на контекста на запитването,
- възможност за формулиране на запитване на естествения език,
- адаптиране на интерфейса в зависимост от потребителя¹⁷⁰ и др.

Адаптивното търсене се реализира чрез използването на онтологии, които представят знанията за информационните обекти и за потребителите. Те дават допълнителни възможности за търсене в системата на електронната библиотека, които накратко могат да бъдат представени в няколко групи:

а) Разширяване и уточняване на запитването (Semantic Query Expansion)

При адаптивно търсене системата може да уточни или разшири запитването, за да оптимизира получените резултати до приемлив размер и да ги ранжира по подходящ начин, вземайки решения за важността на документите на базата на информация за потребителя, като възраст, предпочитания, използвани документи¹⁷¹ и др. Уточняването на запитването може да се приложи в различни фази: преди запитването да бъде изпратено към търсещата система; в процеса на формиране на отговора и за допълнително пренареждане, структуриране или клъстериране на върнатите резултати. При уточняване на запитването всички думи, предоставени от потребителя, се съпоставят и преобразуват в един или повече типове, напр. тема, ключова дума или име на лице; всяка от тях се разширява, като се използват различни типове свойства. Получените резултати се нареждат според интересите на потребителя, отразени в неговия потребителски профил и историята на по-рано проведени информационни запитвания.

Втората цел, която може да бъде постигната чрез семантично търсене, е намирането на ресурси, чиито изрични описания не съвпадат със запитването, но са свързани¹⁷². Например ако търсим документи (или информация) за история на железопътния транспорт, част от резултатите могат да бъдат документи, посветени на „Ракетата” на Джордж Стивънсън*. Това може да се постигне чрез разширяване на информационното запитване чрез включване на понятия/термини, които са семантично свързани.

¹⁷⁰ **Търнина**, Албена. Технологии на семантичния уеб в услуга на електронното обучение. В: *Списание на Софийския университет за електронно обучение*[online]. 2011, 1. [Прегледан 20.06.2018]. Достъпно от: <https://journal.e-center.uni-sofia.bg/f/downloads/2011/Broi%201/A.Turnina.pdf>

¹⁷¹ **Synak**, Marcin, **Kruk**, Sebastian Ryszard. Op. cit.

¹⁷² **Synak**, Marcin, **Kruk**, Sebastian Ryszard. Ibid.

* „Ракета” е първият модерен парен локомотив, конструиран от Джордж Стивънсън. Той извършва първото си пробно пътуване от Манчестър до Ливърпул през 1829 г.

б) Съвместно филтриране (*Collaborative Filtering*)

Потребителят може да бъде информиран какво четат други потребители, какво намират за важно и интересно. Всеки потребител в семантична електронна библиотека е част от мрежа от „приятели”, споделяйки своите интереси. Обикновено, когато даден потребител търси в електронна библиотека, някои статии и материали му се струват по-важни от други. Най-често срещаната практика е да ги маркира (bookmark). Потребителите искат да видят отметките на приятелите си и да използват събраните от тях знания. Социалното семантично сътрудничество при филтриране (*Social Semantic Collaborative Filtering*) е важна услуга на семантичните електронни библиотеки¹⁷³.

в) Фасетно навигиране (*Multifaceted Browsing*)

Използва се за „селектиране” на зададения резултат според индивидуалните предпочитания на потребителя. След получаване на резултат от търсене потребителят може да приложи някои ограничения, (например по година, вид на ресурса, тема и т.н.) и да извърши „търсене в търсене”, като намали наборът от резултати до необходимия обем.

2.2.3. Стандарт SKOS

Системата за опростена организация на знанията (*Simple Knowledge Organization System, SKOS*) е предназначена за представяне и използване на различни системи за организация на знания – класификационни схеми, предметни системи, тезауруси, таксономии или други структурирани нормативни речници в уеб. SKOS е част от множеството стандарти (*W3C Recommendation*) на Консорциума W3C за семантичен уеб, изградени върху RDF и RDFS. Той е резултат от дългогодишна изследователска и практическа дейност по международни проекти като *DESIRE II project* (1997–2000), *LIMBER* (*Language Independent Metadata Browsing of European Resources* 1999–2001), *SWAD-Europe project* (2002–2004 *Semantic Web Advanced Development for Europe*). За първи път е публикуван като препоръка на Консорциума през 2008 г. Основната цел на стандарта е да направи възможно публикуването и използване на системите за организация на знанието като свързани данни¹⁷⁴.

¹⁷³ Kruk, S. R., Decker, S., Gzella, A., Grzonkowski, S., McDaniel, B. Social semantic collaborative filtering for digital libraries. In: *Journal of Digital Information* [online]. Special Issue, 2006. [Viewed 24.05.2018]. Available from: https://www.researchgate.net/profile/Sebastian_Kruk2/publication/242404395_Social_Semantic_Collaborative_Filtering_for_Digital_Libraries/links/5568061908aeced777378f9d.pdf

¹⁷⁴ *SKOS Simple Knowledge Organization System* [online]. [Viewed 28.06.2018]. Available from: <https://www.w3.org/2004/02/skos/>

SKOS е съгласуван с международния стандарт ISO 25964:2011 – International standard for thesauri and interoperability with other vocabularies, въпреки принципната разлика: стандартът регламентира съставянето и използването на тезауруси, независимо дали имат приложение в интернет и машинна обработка или са в традиционна форма, докато SKOS трябва да осигури формалното представяне на всички видове системи за организация на знанието, включително класификационни схеми и много други речници, които не съответстват на определението за тезаурус, съгласно ISO 25964¹⁷⁵.

SKOS се актуализира и съгласува с промените в международните стандарти и се съобразява с действащите модели. Например SKOS напълно отговаря на MADS/RDF (Metadata Authority Description Schema in RDF) – система за организация на знанието, създадена като модел за нормативни данни и нормативни речници, използвани от библиотеки, музеи, архиви и други културни институции. MADS/RDF осигурява средство за записване на данни от MARC 21 формат за нормативни данни” (MARC 21 Format for Authority Data) в RDF за използване в семантични приложения и проекти за свързани данни¹⁷⁶.

SKOS определя класовете и свойствата, които са достатъчни, за да представят общите характеристики на стандартен тезаурус. Системата за опростена организация на знанията фокусира върху понятията, не върху термините, които ги представят за разлика от традиционните системи за организация на знанието. Всяко понятие (concept) в SKOS е определено като RDF ресурс с идентификатор. Понятието има следните RDF свойства (properties):

- има един или повече предпочитани термини за индексирание (при многоезикова система по един на всеки естествен език);
- алтернативни термини или синоними;
- дефиниции и бележки (при многоезикова система със спецификация на всеки от езиците).

Понятията могат да бъдат организирани в йерархия, чрез връзките „род-вид”, „част-цяло” или „висше-низше” в SKOS – broader-narrower или свързани с асоциативни връзки (related). Понятията могат да бъдат групирани в концептуални схеми (ConceptScheme), за да

¹⁷⁵ **Interoperability** with SKOS and other schemas. In: *ISO 25964 – the international standard for thesauri and interoperability with other vocabularies* [online]. [Viewed 20.06.2018]. Available from: <https://www.niso.org/schemas/iso25964#skos>

¹⁷⁶ За повече информация за свързаните библиографски данни (Linked Data) виж: Миланова, Милена. Библиографски данни в семантичния уеб... [онлайн].

се осигури последователно и структурирано представяне на понятията в речника. В концептуална схема мога да влизат всички понятия от речника или само част.

Стандартът SKOS дефинира следните категории: понятия (concepts), етикети (labels), бележки (notations), семантични връзки (semantic relations), свойства (mapping properties), колекции (collections):

Понятия (concepts)

Речникът на SKOS се основава на понятия. Понятие (Concept) е обект, събитие, явление, идея. Може да е отделен екземпляр или клас. В SKOS Concept служи да представи отделните единици (обект, идея и т.н) в организацията на знанието и за създаване на концептуална или организационна структура. Пример за представяне на понятие в SKOS:

```
ex:animals rdf:type skos:Concept.
```

Концептуалната схема на понятията (ConceptScheme) е аналогична, както при речник, тезаурус или друга система за организиране на понятия. SKOS не поставя ограничения за начина на организиране на схемата. Той може да бъде например йерархия на база понятията от най-горно ниво в SKOS – TopConcept.

Етикети и нотация (Labels & Notation)

Всеки етикет (Labels) в SKOS е низ от символи от Unicode, свързан с понятие, т.е. термин, който обозначава понятие. PrefLabel е термин, който може да се чете от човек (по един за всеки език); altLabel се използва за алтернативни термини; hiddenLabel може да се използва за низове, които са полезни за свързване, но не са предназначени за четене от хората. Примери за етикети:

```
ex:animals rdf:type skos:Concept;  
skos:prefLabel "animals"@en;  
skos:altLabel "creatures"@en;  
skos:prefLabel "animaux"@fr;  
skos:altLabel "créatures"@fr.
```

Нотацията (notation) е подобна на етикет, но низът е „тип данни”, напр. като дата. Типът на данните може да бъде дефиниран според необходимостта. Нотациите са полезни за класификационните кодове и други низове, които не могат да бъдат разпознати като думи, напр. индекс 512 от УДК, представен като нотация:

```
ex:udc512 skos:prefLabel "Algebra"@en;  
skos:notation "512"^^ex:UDCNotation.
```

Бележки (Documentation)

Бележките (или документацията) са предназначени да предоставят основна информация за понятията, напр. бележката Definition съдържа информация за предвиденото значение на понятието:

```
ex:documentation skos:definition
"the process of storing and retrieving information
in all fields of knowledge"@en.
```

Семантичните връзки (Semantic relations)

Елементите в категорията “Семантични връзки” в SKOS служат за оказване на взаимовръзките между понятията в една схема.

Related създава асоциативна връзка между две понятия, когато няма йерархична връзка между тях, напр.:

```
ex:birds rdf:type skos:Concept;
skos:prefLabel "birds"@en;
skos:related ex:ornithology.
```

Broader (широко понятие) и Narrower (тясно понятие) се използват за изразяване на пряка йерархична връзка („род-вид”, „част-цяло”, „висше-нисше”) между две понятия, напр.:

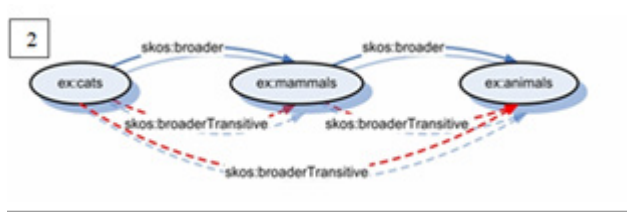
```
ex:animals rdf:type skos:Concept;
skos:prefLabel "animals"@en;
skos:narrower ex:mammals.
```

```
ex:mammals rdf:type skos:Concept;
skos:prefLabel "mammals"@en;
skos:broader ex:animals.
```

Чрез преки йерархични връзки обаче не могат да се изразят транзитивни отношения (виж пример 1).



За тази цел служат транзитивните йерархични връзки, които се изразяват чрез `BroaderTransitive` и `narrowerTransitive` (виж пример 2).



Чрез транзитивните йерархични връзки могат да се изведат следните 2 твърдения:

`ex:cats skos:broaderTransitive ex:mammals.`

`ex:mammals skos:broaderTransitive ex:animals.`

и

`ex:cats skos:broaderTransitive ex:animals.`

Картографиране (Mapping Properties)

На всяко понятие в SKOS е присвоен URI. Това осигурява възможност за съпоставяне (картографиране) на понятия от различни схеми. Картографирането е възможността да се посочи, че две понятия от различни схеми имат сравними значения. Картографирането има съществено значение за информационно-търсещи системи, които използват едновременно няколко системи за организация на знанието и трябва да бъдат съгласувани семантично.

Свойствата `Match`, `broadMatch`, `narrowMatch` др. са предназначени да изразят съответствия, връзки между понятия от една концептуална схема с понятия от друга схема. `RelatedMatch` свързва асоциативно свързани понятия. Когато понятията са толкова тясно свързани, че обикновено могат да бъдат използвани като взаимнозаменяеми, връзката се изразява с `exactMatch`. Ако само в някои случаи могат да бъдат взаимнозаменяеми, се използва `ClosMatch`.

Колекция понятия (Concept collections)

SKOS дава възможност да се определят смислови групи (collections) от понятия. Такива групи се правят обикновено в тезаурси, както в следния пример:

milk

< milk by source animal >

Cow milk

Goat milk

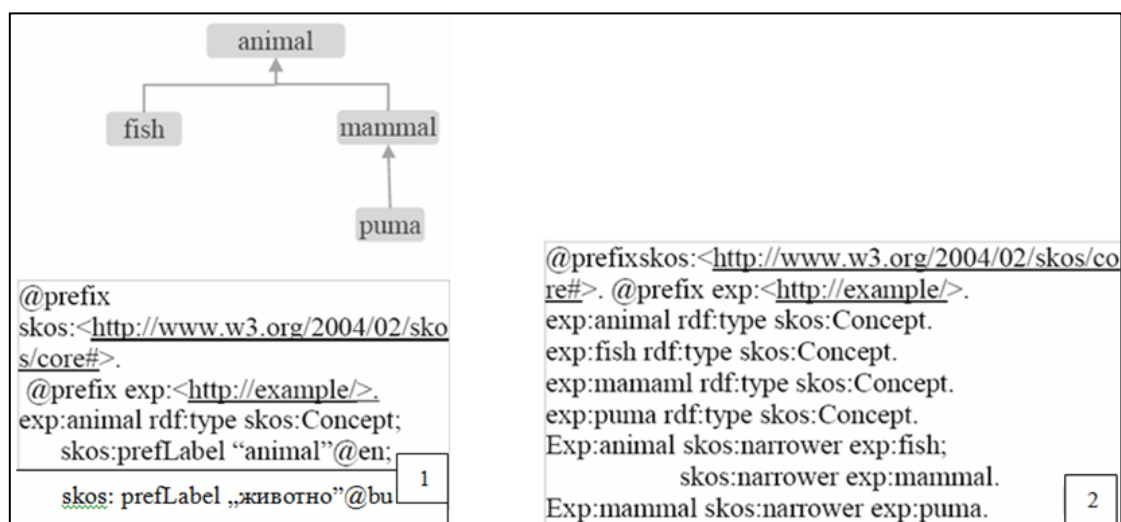
Buffalo milk.

Примерът, представен в SKOS описание:

```
_:b0 rdf:type skos:Collection;
      skos:prefLabel "milk by source animal"@en;
      skos:member ex:cowMilk;
      skos:member ex:goatMilk;
      skos:member ex:buffaloMilk.
```

На фиг. 16 представяме пример на запис в SKOS на понятието „animal”, което има два предпочитани термина, съответно за два езика български и английски (1); и на (2) са изразени преките йерархични връзки с други понятия – „fish”, „mammal”, „puma”.

Фигура 16. SKOS описание.



Приложението на SKOS с времето се разширява. Някои от най-популярните в библиотечно-информационната област системи за организация на знания са мигрирали във формат SKOS и са достъпни в публичното пространство (табл. 5).

Таблица 5. Системи за организация на знанията в SKOS¹⁷⁷.

Система за организация на знанието	Формален език	Уеб адрес	Публикации
Предметни рубрики на Конгресна библиотека	SKOS	http://id.loc.gov/authorities/subjects.html	Summers, E., Isaac, A., Redding, C., Krech, D. LCSH, SKOS and Linked Data. In: <i>Proceedings of the 2008 International Conference on</i>

¹⁷⁷ Източник: **SKOS/Datasets**. In: *W3C Semantic Web* [online]. [Viewed 20.06.2018]. Available from: <https://www.w3.org/2001/sw/wiki/SKOS/Datasets>

Library of Congress Subject Headings, (LCSH)			<i>Dublin Core and Metadata Applications</i> , 2008, p. 25–33.
Математическа предметна класификация Mathematics Subject Classification (MSC)	SKOS	http://msc2010.org/mscwork/	Lange, C., Ion, P., Dimou, A. Bringing Mathematics to the Web of Data: The Case of the Mathematics Subject Classification. In: <i>The Semantic Web: Research and Applications. 9th Extended Semantic Web Conference, ESWC 2012, Heraklion, Crete, Greece, May 27-31, 2012</i> . Berlin: Springer, 2012, p. 763–777.
Компютърна класификационна система Computing Classification System (ACM)	SKOS	http://www.acm.org/about/class/class/2012	Kohl, Christian. ACM Computing Classification System (CCS). Available from: http://taxonomies.labs.crossref.org/?p=109
Медицински предметни рубрики MeSH Medical Subject Headings	SKOS	https://finto.fi/mesh/en/	Assem, Mark van et al. <i>A Method to Convert Thesauri to SKOS</i> . Available from: http://thesauri.cs.vu.nl/eswc06/
Тезаурус AGROVOC	SKOS-XL	http://aims.fao.org/standards/agrovoc	Caracciolo, C. et al. The AGROVOC Linked Dataset. In: <i>Semantic Web Journal</i> , Vol. 4, 2013, p. 341–348. Available from: http://www.semantic-web-journal.net/system/files/swj274_1.pdf
Многоезичен тезаурус на Европейския съюз EuroVoc	SKOS SKOS-XL	http://data.europa.eu/euodp/bg/data/dataset/eurovoc	Release notes – Bulgarian EuroVoc version 4.7. PDF. Available from: http://eurovoc.europa.eu/drupal/sites/all/files/Release_notes_MDR_NAL_publication_eurovoc_20171215-0_bul.pdf
Виртуален международен	SKOS	http://viaf.org	Linked Data and VIAF . In: EMEA 2011. Available from:

нормативен архив Virtual International Authority File (VIAF)			https://www.oclc.org/content/dam/research/presentations/hickey/20110302-EMEARC.pdf
Таксономия на НАСА (NASA taxonomy)	SKOS	https://ntrs.nasa.gov/search.jsp?R=20070038083	NASA Taxonomy: A Way to Unify the NASA Information Space ISKO Singapore. Available from: www.iskosg.org/uploads/3/8/2/6/38261647/nasa_taxonomy.pdf

2.3. Създаване на семантични електронни библиотеки

Семантична електронна библиотека използва технологии на семантичната мрежа и средства на социалния уеб за подобряване на търсенето^{178,179,180} и осигуряване на оперативна съвместимост с други системи (Уеб 3.0)¹⁸¹.

Семантичните услуги подобряват функциите за търсене и навигиране и осигуряват възможност за семантична съвместимост между различни системи¹⁸², докато социалните услуги позволяват на хората да взаимодействат и да се възползват от опита на другите. Съчетани, тези услуги подобряват споделянето на знания¹⁸³.

Внедряването на технологии на семантичния уеб в електронните библиотеки е стъпка към създаването на интелигентни библиотечни системи, които биха предложили по-добра информация, основана на семантични знания за понятията и за потребителските интереси и навици.

Най-съвременните електронни библиотеки трябва да отговарят на основни изисквания и очаквания на потребителите:

¹⁷⁸ **Semantic** Digital Library. In: *Dictionary information science and technology* [online]. [Viewed 24.05.2018]. Available from: <https://www.igi-global.com/dictionary/?p=Semantic+Digital+Library>

¹⁷⁹ **Tella**, Adeyinka, **Okojie**, Victoria, **Olaniyi**, O. T. *Social Bookmarking Tools and Digital Libraries. Key Terms in this Chapter* [online]. [Viewed 24.05.2018]. Available from: <https://www.igi-global.com/chapter/social-bookmarking-tools-and-digital-libraries/188559>

¹⁸⁰ **Kruk**, S. R., **Woroniccki**, T., **Gzella**, Adam, **Dabrowski**, Maciej, **McDaniel**, Bill. *The anatomy of a Social Semantic Digital Library?* [online]. [Viewed 24.05.2018]. Available from: https://www.researchgate.net/publication/267718960_The_anatomy_of_a_Social_Semantic_Digital_Library

¹⁸¹ *Semantic Web Technologies for Digital Libraries JeromeDL – A Case Study* [online]. [Viewed 24.05.2018]. Available from: <https://sites.google.com/a/isim.net.in/nikesh/semantic-digital-libraries>

¹⁸² **The Semantic Web Vision**. In: *A semantic web primer* [online]. 3rd ed. Cambridge: MIT Press, 2012, p. 1–3. [Viewed 28.06.2018]. Available

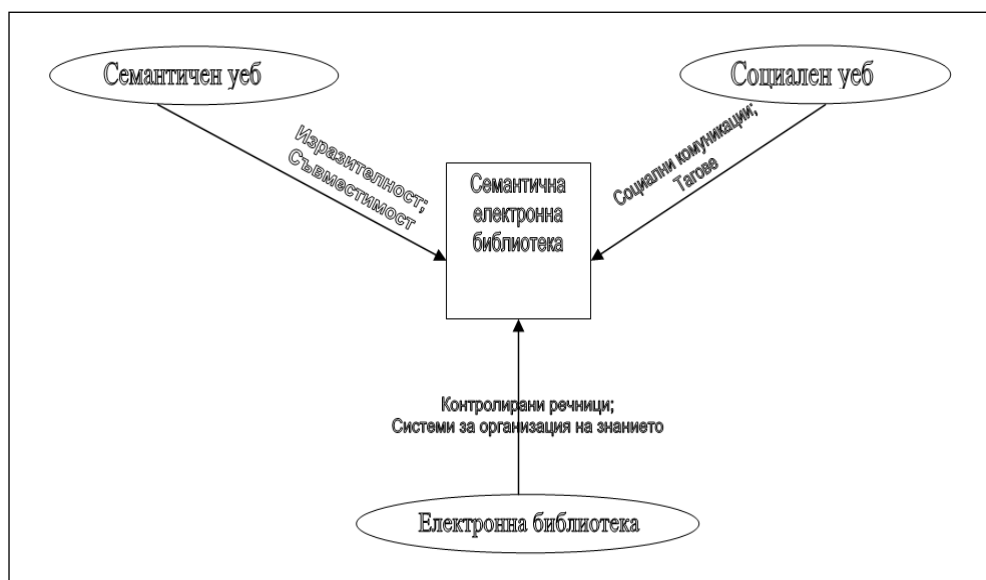
from: [http://prof.mau.ac.ir/images/Uploaded_files/A%20Semantic%20Web%20Primer-The%20MIT%20Press%20\(2012\)\[7460174\].PDF](http://prof.mau.ac.ir/images/Uploaded_files/A%20Semantic%20Web%20Primer-The%20MIT%20Press%20(2012)[7460174].PDF)

¹⁸³ **Kruk**, S. R., **Decker**, S., **Gzella**, A., **Grzonkowski**, S., **McDaniel**, B. Op. cit.

- да осигуряват достъп до множество разпределени и разнородни документи, информация и знания, интегрирани в единно информационно пространство;
- да позволяват възможност за взаимодействие между потребителите;
- да предлагат разширени възможности за търсене.

Семантичната електронна библиотека се изгражда като се базира на системи за организация на знанията, използвани в библиотеките; онтологии за изразяване на семантика и за семантична съвместимост; решения на Социалния уеб за взаимодействие между потребителите. На фиг. 17 са представени основните технологии, които накратко ще разгледаме като изграждащи компоненти на семантична електронна библиотека.

Фигура 17. Семантична електронна библиотека

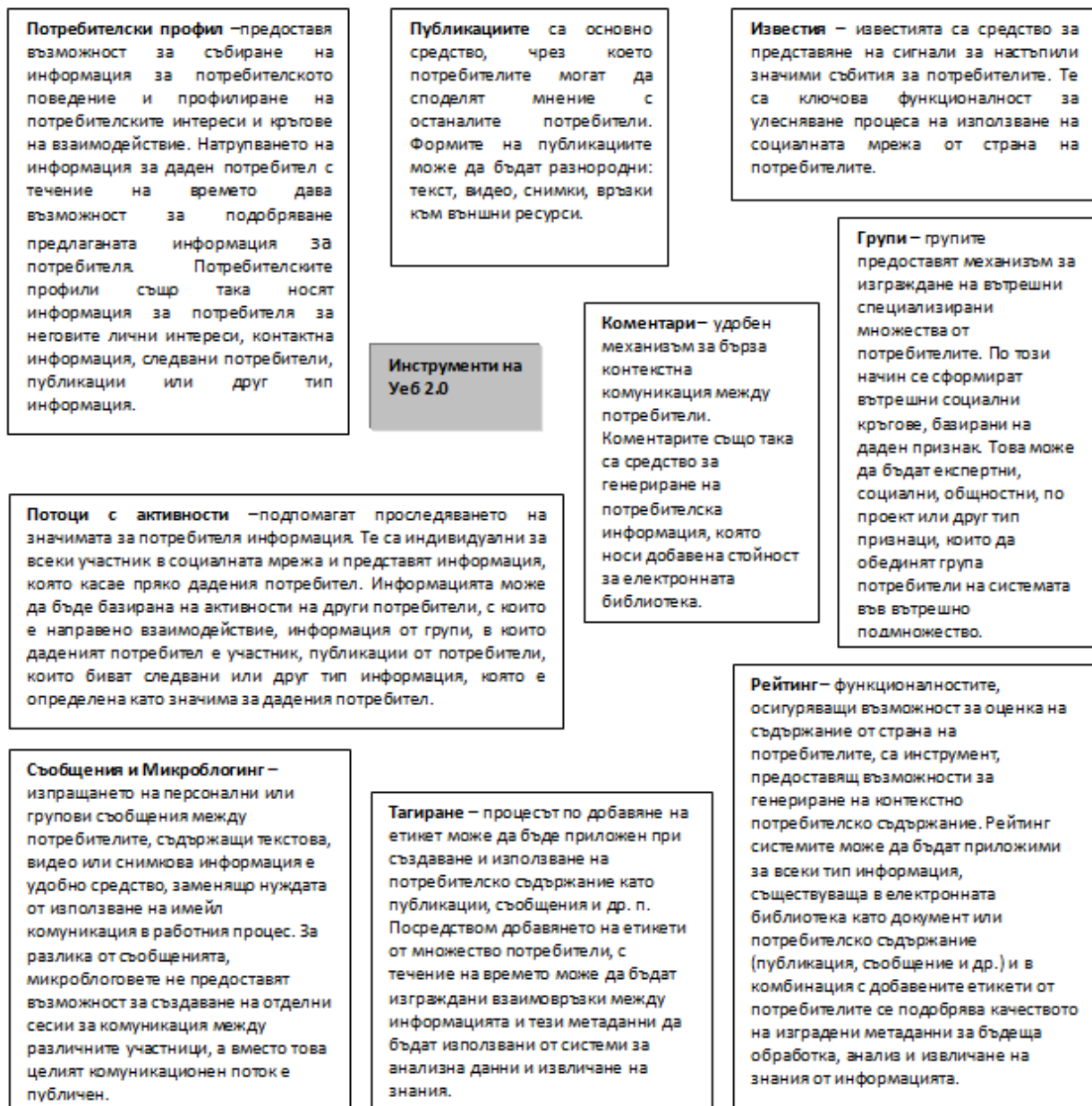


Първият компонент на тематична електронна библиотека са системите за организация на знанията и техните речници, представени като формализирани онтологии за използване в уеб, които вече бяха разгледани в настоящата глава, т. 2.1.3 и 2.2.3.

Вторият компонент са средства, инструменти и технологии за социално взаимодействие и изграждане на общности, характерни за Уеб 2.0¹⁸⁴: потребителски профил, групи, коментари, маркиране (тагиране) и др. Те позволяват на потребителите да взаимодействат и участват в процесите по аотирането на документите и по обмена на знания, което повишава функционалността на електронната библиотека (фиг. 18).

¹⁸⁴ Георгиев, Веселин. Насоки за развитие на корпоративните комуникации чрез приложение на средства от социалните мрежи. В: *Икономически и социални алтернативи*[online]. 2014, 3, с.127–137 [Прегледан 16.06.2018]. Достъпно от: <https://www.unwe.bg/alternativi/bg/journalissues/article/5881>

Фигура 18. Инструменти на уеб 2.0



Третият компонент на семантичната електронна библиотека се технологиите на семантичния уеб, представени в Глава 2, т. 2.2.1.

Реализирането на семантични електронни библиотеки се основава на следните важни задачи:

- осигуряване на висока степен на оперативна съвместимост между информационни системи (не само с електронни библиотеки), чрез използване на RDF;
- възможност за интегриране на информация и метаданни от различни източници: библиографски записи, контролирани речници, профили на потребители и потребителски отметки (bookmarks) и др.;

– търсене и навигиране, основано на онтологии чрез: а) информационни езици, използвани от библиотеките, б) търсене на естествен език и в) допълнителни възможности за търсене – анотации и етикетите, създавани от потребители;

– персонализирано търсене, основано на потребителския профил (отчитащо индивидуалните особености и потребности) и предлагане на препоръки (документи по интересите на потребителя);

– управление на потребители, управление на колекции, управление на онтологични инструменти;

– използване на Peer-to-Peer и Grid* инфраструктура¹⁸⁵.

Като пример за семантична електронна библиотека в литературни източници се разглежда JeromeDL. Тя е създадена като прототип на семантична електронна библиотека по проект на Библиотеката на Техническия университет в Гданск и DERI.International¹⁸⁶. В JeromeDL технологиите на семантичния уеб се използват съвместно с инструменти и средства на Социалната мрежа¹⁸⁷. Индексирането на съдържанието на документите се извършва чрез традиционните за библиотеките системи за организация на знанието – контролирани речници, класификационни схеми, тезауруси и пр. Освен това се прилагат и средства като социално маркиране, потребителски анотации и използване на фолксономии.

Семантичните описания на информационните обекти е на база специално съставена библиографска онтология – MarcOnt Ontology. Благодарение на онтологията описанията могат да бъдат конвертирани в различни формати MARC, Dublin Core и BibTeX. Функцията

* Grid инфраструктурата се определя като мрежа от разпределени компютърни ресурси, достъпни чрез локални или глобални мрежи, които се визуализират пред крайния потребител като огромна виртуална система. Източник: Дигитални библиотеки с мултимедийно съдържание и приложения в българското културно наследство [онлайн]. София: БАН, 2005. PDF. с. 22. [Прегледан 16.06.2018]. Достъпно от:

http://mdl.cc.bas.bg/Digital_libraries_with_multimedia_content_and_applications_in_Bulgarian_cultural_heritage.pdf

¹⁸⁵ Biagetti, Maria Teresa. Digital libraries and semantic searching. In: *Knowledge organization in the 21st century: between historical patterns and future prospects* [online]. Proceedings of the Thirteenth International ISKO Conference, 19–22 May 2014, Kraków, Poland. Würzburg: Ergon Verlag, 2014, p. 486–492. [Viewed 16.06.2018]. Available from:

https://www.researchgate.net/publication/263924292_Digital_Libraries_and_semantic_searching_2014; Kruk, Sebastian Ryszard, McDaniel, Bill. Goals of Semantic Digital Libraries. In: *Semantic Digital Libraries* [online]. Berlin Heidelberg: Springer, 71–76; Kruk, Sebastian Ryszard, Decker, Stefan, Zieborak, Lech. JeromeDL – Adding Semantic Web Technologies...; Kruk, Sebastian Ryszard, Woroniecki, Tomasz, Gzella, Adam, Dąbrowski, Maciej. JeromeDL: The Social Semantic Digital Library. In *Semantic Digital Libraries* [online]. Berlin Heidelberg: Springer, 2009, p. 139–150.

¹⁸⁶ JeromeDL e-library with semantics [online]. [Viewed 24.06.2018]. Available from:

http://www.isin.ie/assets/13/1E513B27-00B2-64D5-516A8120E911BA4E_document1/1__jeromedl.pdf

¹⁸⁷ Kruk, Sebastian Ryszard, Woroniecki, Tomasz, Gzella, Adam, Dąbrowski, Maciej. JeromeDL – a Semantic Digital Library [online]. [Viewed 24.06.2018]. Available from:

https://www.researchgate.net/publication/221466019_JeromeDL_-_a_Semantic_Digital_Library

за търсене се реализира по различни методи и средства: пълнотекстов анализ, търсене по библиографски данни (автор, заглавие предмет, и т.н.), разширяване и филтриране на потребителското запитване¹⁸⁸.

JeromeDL описва всеки документ като използва различен тип метаданни: метаданни, които описват структурата на ресурса (напр. `jerome:cover`, `jerome:abstract`, `jerome:hasChapter`); библиографски метаданни (напр. `marcont:hasTitles`, `marcont:hasCreator`, `marcont:hasDomain`, `marcont:hasKeyword`). (Фиг. 19):

Фигура 19. Пример от JeromeDL¹⁸⁹

The screenshot displays a digital library interface for a document titled "Building a Heterogeneous Network of Digital Libraries on the Semantic Web". The interface is structured as follows:

- Title Section:** Displays the title "Building a Heterogeneous Network of Digital Libraries on the Semantic Web" and authors "Sebastian Ryszard Kruk, Mariusz Cygan, Piotr Piotrowski, Krystian Samulski". It includes a mapping: `marcont:hasTitles -> marcont:titleValue`.
- Abstract Section:** Shows a text preview of the abstract and a mapping: `jeromedl:abstract`.
- Chapter Section:** Indicates the document is an "article" and shows a chapter mapping: `jeromedl:hasChapter`.
- Domains Section:** Lists the domain "Systems" and a mapping: `marcont:hasDomain`.
- Keywords Section:** Lists keywords: "network, library, heterogeneous, digital, digital libraries" and a mapping: `marcont:hasKeyword`.
- Additional description Section:** Lists properties: "presentedAt: Semantics2006", "cites: http://library.der1.ee/resource/5e68d6a6", and "year: 2006". It includes a mapping: `marcont:x*`.
- Cover Section:** Shows a placeholder for the document cover with the mapping: `jeromedl:cover`.

Архитектурата на JeromeDL е изградена от три слоя. Всеки слой обогатява информацията за документа и осигурява допълнителни възможности за търсене. Най-долният слой отговаря за типичните задачи, изисквани от хранилище на цифрови обекти, т.е. следи за физическото им представяне чрез използването на структурна онтология. Вторият слой се основава на библиографска онтология MarcOnt, която прави библиографските описания семантични. Услугите, предоставяни от този архитектурен слой са фокусирани върху съхраняването, доставянето и управлението на метаданни за документите. Най-горният слой в архитектурата на семантичната библиотека JeromeDL

¹⁸⁸ Synak, Marcin, Kruk, Sebastian Ryszard. Op. cit.

¹⁸⁹ Kruk, Sebastian Ryszard, Woroniecki, Tomasz, Gzella, Adam, Dabrowski, Maciej. Op. cit.

използва предимствата от ангажиране на потребителската общност за аотиране и съвместно филтриране¹⁹⁰.

През последните 10–15 години са проведени и други резултатни изследвания и проекти в световен мащаб за внедряване на семантични технологии за създаване на оптимизирани електронни библиотеки:

– SIMILE (Semantic Interoperability of Metadata and Information in unLike Environments, 2003) е съвместен изследователски проект, ръководен от Консорциума на W3 и Масачузетския технологичен институт, финансиран от Фондация Андрю У. Мелън (Andrew W. Mellon Foundation). Екипът работи по разширяването и развиването на възможностите на една от най-популярните програмни системи за организация и управление на електронни ресурси – DSpace. Ключови аспекти на този проект са подобряването на взаимодействието между хранилищата на данни и взаимодействието между схеми, речници, онтологии за повишаване качеството на метаданните и съответно услугите¹⁹¹.

– Проект DELOS, който препоръчва внедряването на технологиите P2P, Grid, SOA в електронни библиотеки.

– Проект BRICKS (Building Resources for Integrated Cultural Knowledge Services, 2004-2007) цели разработването и поддържането на ориентирано към потребители и услуги пространство за обмен на знания в областта на културното наследство¹⁹².

Изводи:

1. Съществуващите обективни затруднения в търсенето и обмена на информация в електронна среда се предопределят от ограниченията на информационните системи, които не могат да „разбират“ естествения език, използван от потребителите. Разработените логически езици за комуникация между машина и машина и между машина и човек не постигат разбиране на смисъла на потребителското запитване. Съвременните усъвършенствани информационни системи се изграждат на базата на най-новите достижения на науката – изкуствен интелект, компютърна лингвистика, организация на

¹⁹⁰ Ibid.

¹⁹¹ *SIMILE Overview: Semantic Interoperability of Metadata and Information in unLike Environments* [online]. [Viewed 20.06.2018]. Available from: <http://web.mit.edu/dspace-dev/www/simile/resources/overview.html>

¹⁹² *BRICKS Project ID: 507457*. In: *CORDIS* [online] Funded under: FP6-IST. [Viewed 20.06.2018]. Available from: https://cordis.europa.eu/project/rcn/71256_en.html

знанията. Целта е взаимодействие с потребителя на естествен език, създаването на семантични връзки в системата за по-ефективно търсене и извличане на информация.

2. Приносът на изкуствения интелект в развитието на информационните системи, вкл. и на електронни библиотеки, касае създаването на модели за формализирано представяне и използване на знания и на методи за тяхната машинна обработка. Компютърната лингвистика подпомага решаването на редица задачи, свързани с автоматичната обработка на естествения език, като изгражда лексикални семантични ресурси. Те могат да бъдат използвани за концептуално индексирание и търсене в електронните библиотеки. Организацията на знанието е тясно свързана с библиотечно-информационните науки, като в тесния си смисъл има най-изявена връзка с класификациите на информационните ресурси, тяхното описание и индексирание. Системите за организация на знания (като напр. семантични мрежи, онтологии, метаданни, тезауруси и мн.др.) улесняват управлението и търсенето на информация и знания от хора и машини. Макар че имат различна семантична сила, те широко се използват за изграждане на лингвистичната основа на електронните библиотеки. Важен аспект е постигането на семантична съвместимост между електронните библиотеки на основата на онтологии с различни функции.

3. Разработваният Уеб 3.0, определен като семантичен, е призван да разреши проблема с автоматичната обработка на големи обеми от информация, семантичното търсене, както и улесняването на достъпа на основа на свързаните данни. Той ще осигури обща среда, позволяваща информацията да бъде обработвана съвсемстно и многократно от потребители и софтуерни приложения. Семантичният уеб се разработва на базата на уеб стандарти, метаезици, онтологии и логически правила.

Глава 3. ЛИНГВИСТИЧНО ОСИГУРЯВАНЕ НА ЕЛЕКТРОННА БИБЛИОТЕКА

3.1. Лингвистичното осигуряване като компонент на електронната библиотека

В специализираната литература словосъчетанието „лингвистично осигуряване” се появява през 60-те години на ХХ в. и се използва наред с понятието информационни езици. Както в началото, така и днес „лингвистично осигуряване” не носи нормативен характер и не е фиксирано като точно определение в определен стандарт, но изначално се използва като събирателно понятие, което обхваща езиците за индексирание на документи и информационни запитвания, предназначени за описание, идентифициране, структуриране, систематизиране и търсене на документи и информация.

Терминът „лингвистично осигуряване” се определя по пътя на причисляването на едни или други изкуствени езици, методи и средства за обработка на информация, обусловени от спецификата на конкретната автоматизирана информационна система. Най-общо лингвистично осигуряване означава създаване на комплекс от езици за работата на информационна система. Като компонент на информационната система лингвистичното осигуряване има следните функции: представяне на документите в системата, описание на характеристиките на документите за адекватното им интерпретиране от системата и осигуряване на взаимодействие между системата и потребителя¹⁹³.

Според изследователи¹⁹⁴, въпреки натрупаното количество изследвания по това научно направление, липсва еднозначно тълкуване на понятието „лингвистично осигуряване”. Трудността при дефиниране на термина свързваме с две основни причини:

Първата от тях: „лингвистично осигуряване” се използва с многообразни обекти и съответно различни области на прилагане:

- лингвистично осигуряване на различни типове и видове информационни системи - автоматизирани, библиотечни, системи за управление, системи за проектиране и т.н.;
- лингвистично осигуряване на конкретни организации или системи;
- лингвистично осигуряване на конкретни технологии или процеси, например: търсене на информация, проектиране, комплектуване и др.;
- лингвистично осигуряване на различни информационни ресурси, например: сайтове, архиви, бази или банки от данни т.н.

¹⁹³ **Коголовский**, М. Р. *Перспективные технологии информационных систем*. Москва: ДМК, 2003, с. 13–14.

¹⁹⁴ **Гендина**, Н. И. *Лингвистические средства библиотечно-информационных технологий*. Санкт-Петербург: Профессия, 2015, с. 14–15.

Така на базата на терминоелемента „лингвистично осигуряване” се образуват многочислени термини с различен смисъл: например „лингвистично осигуряване на автоматизираните библиотечно-информационни системи”, „лингвистично осигуряване на електронните библиотеки”, „лингвистично осигуряване на системите за автоматизирано проектиране”. Този ред може да бъде продължен, но и от посочените примери става ясно, че обектите на лингвистично осигуряване могат да бъдат различни. Вследствие на това терминът „лингвистично осигуряване” има широк обхват на използване в различни научни дисциплини и предметни области – теорията на информационно-търсещите системи, информатиката, библиотечното дело, лингвистиката и др. При това „лингвистично осигуряване” може да се използва като абстрактно понятие и като конкретно понятие, свързано с реален обект. Следователно обемът и съдържанието на понятието „лингвистично осигуряване” зависи от обекта и сферата на приложение.

Втората причина, която затруднява дефинирането на „лингвистично осигуряване” е многозначността на думата „език”. В областта на информатиката и компютърните технологии „език” се използва за обозначаване на различни средства за описание на данни и алгоритми за решаване на задачи. В процеса на развитие на компютърните и информационните системи съставът на езиковите средства се променя и разширява. Например като средства за лингвистично осигуряване могат да се разглеждат езиците за описание на данни (Data Definition Language, DDL) и за манипулация на данните (Data Manipulation Language, DML) в бази от данни, или маркиращите езици, езици за програмиране и много др.

В литературата се представят различни подходи при разглеждане на понятието „лингвистичната осигуреност” на автоматизираните информационни системи. Най-разпространен, според изследователите е „класическият подход”, при който лингвистичната осигуреност се разглежда като „комплекс от информационно-търсещи езици, преди всичко – класификационните и словесните”¹⁹⁵. Този подход води началото си от класическите трудове по информатика¹⁹⁶ и е типичен за системите за научно-техническа информация. С незначителни изменения този подход е приет и в теорията на автоматизираните библиотечно-информационни системи. Като към лингвистичното осигуряване на тези системи се причислява и езикът на библиографското описание.

Съществува подход, развиван от специалистите по приложна и компютърна лингвистика, според който лингвистичното осигуряване на информационните системи

¹⁹⁵ Антопольский, А. Б. *Цит съч.*, с. 50.

¹⁹⁶ Михайлов, А. И., Черный, А. И., Гиляревский, Р. С. *Основы информатики*. Москва: Наука, 1968, с. 757.

включва преди всичко приложенията или способите и методите за автоматично анализиране на текстове на естествен език, за извличане на информация, за машинен превод и др. п.

Като по-общ се определя „семиотичният подход”, който изхожда от възгледа за знаковата природа на езика. Езиковото осигуряване, според този подход, представлява средствата за представяне на данни и за интерпретиране на тези данни и включва, например средства за кодиране на символи¹⁹⁷, формати за представяне на данни и др.п.

В състава на лингвистичната осигуреност на автоматизираните информационни системи, според програмистите могат да бъдат включени различни езикови средства, не само средствата за описание на данните и средствата за манипулиране с тях, но и инструменталните средства за програмиране.

Процесът на търсене в автоматизираните информационни системи се поддържа с помощта на специализирани информационни технологии и инструменти, които могат да бъдат разделени в две групи: а) алгоритми и програми за автоматична обработка на текстова информация, в основата, на които стоят конкретни алгоритми основани на морфологичен и синтактичен анализ и б) лингвистични банки от данни, които съдържат речници и индекси за търсене и съответен софтуер. Практически, реализирането на лингвистичната осигуреност на автоматизирани информационни системи е създаването и функционирането на лингвистична банка от данни. Тази дейност е свързана, както с информатиката, така и с компютърната лексикография и може да бъде обект на отделно изследване.

В дисертационното си изследване ще се придържаме към класическия подход и ще бъдат изследвани информационните езици за представяне на съдържанието и характеристиките на информационните обекти в електронните библиотеки и осъществяване на търсене. Следователно, без определението да претендира за теоретична „чистота” и точност, под „лингвистична осигуреност” на електронна библиотека разбираме комплекса от информационните езици „предназначени за индексирание на документи и информационни запитвания с цел съхраняване и търсене на нужната информация”¹⁹⁸.

Лингвистичната осигуреност на електронната библиотека включва различни информационни езици, които имат различно предназначение и роля:

¹⁹⁷ Антопольский, А. Б. Цит. съч., с. 50.

¹⁹⁸ Информационно-търсещ език. В: Гергова, Ани, състав. *Енциклопедия Българска книга*. София: Пенсофт, 2004, с. 207.

1. Цялостно представяне на информационните обекти (документите)* – език на библиографското описание (описателни метаданни);
2. Отразяване предмета (темата) на информационните обекти - класификационни езици;
3. Представяне съдържанието на информационните обекти чрез думи от естествения език, съдържащи се в самия обект (или свързани с тях) – посткоординатни езици.

3.2. Информационните езици в състава на лингвистичното осигуряване на електронната библиотека

Преди да представим конкретния модел на лингвистично осигуряване на електронна библиотека, разработен в настоящото дисертационно изследване, бихме искали да направим един обстоен преглед на различните типове и видове информационни езици и техните специфични характеристики. Считаме това за необходимо, за да може предложението от нас модел да бъде осмислен в пълнота.

Най-важният структурен елемент на информационните системи, от които пряко зависят показателите за ефективност на търсенето (пълнота и точност), са информационните езици¹⁹⁹. Редица публикации от 70–80-те години на XX в. доказват необходимостта от създаването и използването на специални изкуствени езици за провеждане на информационно търсене като аргументирано посочват „недостатъците“ на естествения език, възпрепятстващи ефективното търсене²⁰⁰. През този период публикуват Д. Солтън, Ф. Ланкастър, Ч. Медоу, А. Черний, В. И. Москович, както и други учени и изследователи, които разработват теорията на автоматизираното търсене и теорията на информацията и в частност на информационните езици.

На български език постиженията на теорията на информацията и развитието на информационните езици през периода 60–90-те години са представени от Ст. Чавдаров²⁰¹,

* В тази глава под „информационни обекти“ имаме предвид първични информационни обекти или по-традиционното понятие документи. Информационни обекти и документи използваме като синоними, за да избегнем повторенията.

¹⁹⁹ Гендина, Н. И. Цит. съч., с. 36.

²⁰⁰ Williams, Robert V. *Bibliography of the history of information science and technology, 1900–2012*[online]. 8th Ed. [Viewed 15.04.2018]. Available from: <http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=ACF8FFD796125F300BD764CFB616F1F1?doi=10.1.1.676.6682&rep=rep1&type=pdf>

²⁰¹ Чавдаров, Станко. *Теория и методи на научната информация: Обработка, търсене, разпространяване*. София: Наука и изкуство, 1984, 151 с.

Д. Костова и С. Минкова²⁰². В публикуваните от тях учебници за висши училища са описани накратко информационните езици – същност, предназначение и видово разнообразие.

Изследванията на български език, посветени на конкретни информационни езици и тяхното използване при индексирание и информационно търсене също са съвсем ограничени. Българските библиотекарски разполагат с една единствена методика за индексирание, в частност предметизиране, и то свързана с традиционните фишови каталози²⁰³. В средата на 70-те години, когато в България се изграждат системи за научна информация от дескрипторен тип, излизат публикации посветени на методичните и организационните проблеми при съставянето на тезауруси, по-конкретно в областта на обществените науки²⁰⁴.

Изключение правят, особено през последните години, публикациите на изявени научни сътрудници от Националната библиотека „Св. св. Кирил и Методий” (НБКМ) посветени на използването на Универсалната десетична класификация (УДК) в България, и в частност от Националната библиография, както и за развитието на УДК като международна система за организация на знанието в уеб среда²⁰⁵. През 2017 г. е осъществен превод на съкратено издание на таблиците на УДК на български език²⁰⁶. Изданието обичайно е придружено от въвеждаща част, където освен структурните части на класификацията, на основно ниво са представени и технологията на индексирането на документи с УДК и общи правила за индексирание.

²⁰² **Костова**, Десислава, **Минкова**, Соня. *Библиотекознание и теория на научната информация*. София: ВМГИ, 1990, 99 с.

²⁰³ **Атанасова**, В. *Ръководство за предметизация на литературата в библиотеките*. София: НБКМ, 1972, 88 с.

²⁰⁴ **Габровска**, Свободозаря, **Пейчева**, Веселина. Методични и организационни проблеми при съставянето на тезауруси в обществените науки. В: *Съвременни проблеми на научната информация*. София: БАН, 1974., с.29–60.

²⁰⁵ Някои от най-актуалните публикации, свързани с промените в УДК: **Тотоманова**, Антоанета. *Стандартизация на класификационната система в България* [презентация]. Национална кръгла маса „Стандартизацията в библиотеките” 20–21 окт. 2016, София. [Прегледан 15.06.2018 г.]. Достъпно от: https://www.lib.bg/kragli_masi/atotomanova-2016.ppt; **Дипчикова**, Александра, **Тотоманова**, Антоанета. Универсална десетична класификация – съвременни направления в развитието и прилагането им в България. В: *Библиотека*, 2013, № 2–3, с. 99–108.; **Дипчикова**, Александра, **Тотоманова**, Антоанета. УДК – инструмент за оперативна съвместимост на достъпа по съдържание в многоезична мрежова среда [online]. В: *Годишник на Българския информационен Консорциум* [онлайн]. 2013, с. 17–21. [Прегледан 15.06.2018 г.]. Достъпно от: <https://www.bic.bg/files/booklet.pdf>; **Дипчикова**, Александра, **Тотоманова**, Антоанета. *Промените в системите за организация на знанието – предизвикателство към българската библиотечна общност* [презентация]. XXI Национална годишна конференция на ББИА, 10 юни 2011, Благоевград. [Прегледан 15.06.2018 г.]. Достъпно от:

www.lib.bg/konferencii1/nk2011/prezentacii/atotomanova_adipchikova.ppt

²⁰⁶ **Универсална десетична класификация: Съкратено издание**. София: Нац. библ. „Св. св. Кирил и Методий”, 2017, 446 с.

Тъй като липсва цялостно изследване на български език на информационните езици, които се използват в библиотечно-информационните системи намираме за необходимо да им отделим специално внимание като акцентираме върху същността и предназначението им, тяхната структура, типологичното им разнообразие и предназначение. Това ще позволи да открием предимствата и ограниченията на различните информационни езици и ще обоснове подбора на конкретни информационни езици в модела на лингвистично осигуряване на българска електронна библиотека с универсална тематика.

Информационните езици са вид изкуствени езици^{*}, позволяващи да се алгоритмизира процеса по обработката на метайнформацията за информационните обекти. Те получават широко разпространение в сферата на производството, обработката, съхранението и търсенето на информация в различни области и сфери на приложение. Използват се в книгоиздаването и книготърговията, библиотечната, информационната и архивната дейност, документалистиката, деловодството и др. Причина за създаването на информационните езици са „недостатъците“ на естествения език от гледна точка на информационното търсене. „Недостатъците“, всъщност са най-ценните характеристики, за всеки естествен език – богатство, „излишество“ и многозначност^{**}.

Информационните езици са предназначени за търсене и намиране на нужната информация за документи, затова в литературата се срещат и под названието информационно-търсещи езици²⁰⁷. Те служат за съставяне на: 1. формално представяне на документ, т.е. описание на документ, което предполага подбор, анализ, организация и записване на набор от данни; 2. описание на съдържанието на документ или информационно запитване, получено чрез присвояване на термини за индексирание. Описанието на документа, включително и описанието на съдържанието в съвременните информационни системи може да се постигне, както по интелектуален път, така и автоматично чрез програмни средства.

^{*} Изкуственият език е конструиран език, а не такъв развил се по естествен начин в обществото. Той се създава за изпълнението на определени цели.

^{**} Богатството на естествения език се опреля от наличието на синоними, които правят езика по-изразителен. При извършване на информационно търсене, обаче синонимията значително намалява ефективността, като понижава пълнотата на извлечената в резултат на запитването информация; „Излишество“ (редундантност) на естествения език, т.е. наличие в езика на думи, които носят малка смислова натовареност (съюзи, предлози, междуметия), както и на думи, които могат да бъдат пренебрегнати при предаването на съдържанието на документа или информационното запитване; Многозначността – наличието в естествения език на многозначни думи (омоними и полисеми), е признак на смислово богатство и изразителна сила на езика. При информационното търсене обаче полисемията и омонимията играят негативна роля. Те нарушават качеството на информационното търсене като водят до „информационен шум“.

²⁰⁷ **Информационно-търсещ език.** В: Гергова, Ани, състав. *Енциклопедия Българска книга*. София: Пенсофт, 2004, с. 207.

Информационните езици, за да изпълняват своето предназначение, трябва да отговарят на някои специфични изисквания:

- еднозначност – всеки термин изразен на информационен език трябва да има само едно значение. Това изискване предполага отстраняването на такива „явления” в естествения език като полисемия, омонимия, синонимия, свързани с богатството на езика;
- способност да изразяват точно и пълно полезни за търсенето смислови отношения между думите – логически връзки и психологически асоциации;
- разширяемост и гъвкавост на информационния език, възможност за редактиране и попълване на нови понятия, без това да нарушава приетата структура;
- удобство при използване, което предполага стандартно означение, мнемоника и компактност на термините за индексирание;
- да притежават достатъчна семантична сила, т.е. способност на езика да идентифицира понятието, да го отграничи от други и да го опише в достатъчна степен на детайлност.

Информационните езици изпълняват различни функции като компоненти на лингвистичното осигуряване на електронната библиотека:

- търсене – осигурява намирането на документ/и в информационния масив;
- преобразуване – индексирание на изходния документ, с цел последващо търсене;
- структуриране – групиране на еднородни документи по определен признак (категория) и явно изразяване на връзките между отделните групи.
- систематизиране – подреждане на групи от еднородни документи в определен ред;
- идентифициране – еднозначно разграничаване на един документ от други сходни документи.

3.2.1. Структура на информационните езици

Информационните езици са конструирани на базата на естествените езици и притежават, както прилики, така и съществени отличителни характеристики. В структурно отношение всеки език се състои от азбука, речников състав, т.е. лексика и отношения между лексикалните единици, и граматика²⁰⁸.

Структурата на информационните езици ще бъде представена накратко, което ще позволи да открием съществени признаци, по които да бъдат класифицирани с цел

²⁰⁸ Чавдаров, Станко. Цит. съч. с. 73.

подбиране на подходящи езици за модел на лингвистично осигуряване на българска електронна библиотека с универсална тематика.

Азбука на информационните езици.

Азбуката на информационните езици е съвкупността от използвани букви, цифри и символи. Тя може да бъде съставена от букви от латинската или кирилската азбука, арабски и римски цифри, пунктуационни знаци (точка, запетая, двоеточие, кавички, дефис, кръгли и квадратни скоби, апостроф), знаци за аритметични операции (плюс, минус, коса черта, знак за процент), знаци за отношения (по-голямо, по-малко, равно), знаци за конюнкция, дизюнкция и отрицание, специални знаци и символи, например стрелки, двойно двоеточие, квадрат, знак за конгруентност и др.

Азбуката на информационен език може да бъде представена с цифри (кодове), например в УДК, Десетична класификация на Дюи* (ДКД) и др. или с букви, като се основава на азбуката на някой от естествените езици - както е в езика на ключовите думи, на дескрипторите на предметните рубрики, на библиографското описание. Освен това азбуката може да бъде смесена буквено-цифрова, например в Библиотечно-библиографическа класификация (ББК), фасетни класификации и др.

Лексика на информационните езици.

От азбуката се образуват отделни лексикални единици на информационния език, които в съвкупност формират неговата лексика (речников състав). Лексикалната единица представлява последователността от букви, цифри, знаци, приета за обозначаване на едно понятие. Всяка лексикална единица се състои от: обозначение (израз от символи) и съдържание (значение) – понятието, което се представя чрез лексикалната единица. При някои информационни езици, напр. езика на ключовите думи, езика на предметните рубрики, дескрипторния език, обозначението и съдържанието на лексикалната единица съвпадат, тъй като те се основават на азбуката и лексиката на естествен език. При кодовите информационни езици като УДК, ДКД, ББК, Международната патентна класификация** (МПК), Международна класификация на стандартите*** (МКС) и др., „изразът“ не съвпада със значението, тъй като е кодирано, например:

* Dewey Decimal Classification (DDK).

** International Patent Classification (IPC).

*** International Classification for Standards (ICS).

Информационен език	Израз на лексикалната единица	Съдържание на лексикалната единица
УДК	338.48	Туризъм
Език на предметните рубрики	Туризъм	Туризъм

Лексикалните единици във всеки тип информационни езици имат различни наименования: в езика на ключовите думи се наричат ключови думи или унитерми, в дескрипторния език – дескриптори, в езика на предметните рубрики – предметни рубрики, във фасетната класификация – изолати, в УДК – индекси.

Лексиката на информационните езици се формира от термини. Терминът е дума или словосъчетание с ограничена сфера на употреба (наука, техника, изкуство, спорт и т.н.), точно и еднозначно обозначаваща понятие. Термините има точно определение обикновено в ограничена сфера на употреба и се характеризират с качества като: еднозначност (съответствие между термин и понятие), точност, краткост, системност (отразява връзката със свързани понятия), стилистична неутралност, липса на експресивност и др.

Парадигматични отношения.

Между лексикалните единици на информационния език съществуват определени връзки, които се дължат на наличието на смислови и логически отношения между предметите и явленията в реалния свят. Тези връзки имат извънтекстови характер и са независими от ситуацията, в която се използват думите. *Примери, илюстриращи различни смислови и логически връзки:*

- *акварел, темпер и фреска са думи свързани в отношение на съподчинение и влизат в по-общия клас живопис;*
- *между думите широколистни дървета и бреза, дъб, клен съществува отношение „род-вид“;*
- *думите сграда и стена са свързани чрез връзката „цяло-част“;*
- *замърсяване на околната среда и нарушаване на екологичното равновесие са в отношение от причинно-следствен характер;*
- *думата зима извиква устойчива асоциация със думите сняг и студ.*

Парадигматичните отношения са логически, обективно съществуващи отношения между лексикалните единици на информационния език, които не зависят от контекста. Те са обусловени от наличието на нетекстови, а логически връзки между предметите и явленията.

Изследователите разглеждат парадигматичните отношения в две групи – силни и слаби, с оглед ролята, която изпълняват при провеждане на информационно търсене²⁰⁹:

а) Силните парадигматични отношения включват отношенията еквивалентност и йерархия:

– отношението на еквивалентност (съвпадение или синонимия) има пряко въздействие върху качеството на информационното търсене. Допускането на синонимия води до осезаеми загуби в търсенето, като намалява пълнотата на извлечената при запитването информация;

– отношенията на йерархия или подчинение се основава на два вида връзки: „род-вид” (генерични връзки) и „цяло-част” (партитивни връзки);

⇒ „род-вид” е един от най-важните видове връзки между понятията. Родово е подчиняващото понятие, което изразява съществените признаци на клас предмети, представляващи видове на този род. Съответно, видово (или подчинено) се нарича понятие, което показва съществените признаци на даден клас предмети, които са вид на някакъв род;

⇒ „цяло-част” отразява реалното вместиране на един обект (понятие) в състава на друг. При това подчиняващото понятие се съотнася с подчинените в резултат на разпадането на цялото на части.

Значението на силните парадигматични отношения с йерархичен характер („род-вид” и „цяло-част”) за информационното търсене е много важно. Родово-видовите връзки, експлицитно представени в езика, позволяват коригиране на стратегията за търсене, в зависимост от получените резултати. В случай на незадоволителен резултат, използването на родови термини, ще позволи разширяване на обхвата на търсенето, като повиши пълнотата на извлечената информация. Ако в отговор на търсенето се получи огромен обем от информация (информационен шум), то в предписанието за търсене трябва да се прецизира с видови термини, което стеснява параметрите на търсенето и подобрява точността. По същия начин стратегията за търсене може да бъде коригирана и като се вземат под внимание отношенията „цяло-част”, тъй като е много вероятно за потребителя, който се интересува от предмет, система, механизъм, апарат, процес и т.н. като цяло, информацията за съставни части, елементи, детайли, операции да се окаже също много полезна.

б) Слабите парадигматични отношения отразяват логически отношения между понятия или психологически асоциации. Психологическите асоциации възникват в

²⁰⁹ Гендина, Н.И. Цит. съч., с. 55.

съзнанието на човека, когато представата за един предмет (процес или явление) извиква представа за друг/и предмети, свързани чрез някакъв съществен признак. Към категорията на слабите парадигматични отношения се отнасят причинно-следствените отношения, отношенията на близост, контраст, делимост:

– отношението на контраст е резултат от свързването на понятия не по признак за сходство, а по признак, който ги различава или противопоставя. Отношението на контраст се изразява в лексиката на информационния език чрез антоними, напр. Блокиране – Деблокиране;

– отношенията от типа „причина-следствие” характеризират връзка, при която едно явление, действие (причина) предизвиква, определя, променя или създава друго явление. Това отношение свързва лексикалните единици, които обозначават съответните причина и следствие, напр. Авария – Травма;

– отношението на близост (свързаност) отразява асоциативна връзка между предметите или явленията в съзнанието на хората, които се възприемат в непосредствена близост, напр. Електрически ток – Напрежение;

– отношението на делимост (кратност) е специфичен вид връзка, която съществува при категориите мерни единици, тъй като при тях не са приложими други видове отношения.

Синтагматични отношения.

Синтагматичните отношения в информационния език възникват в процеса на използване на лексикални единици, при индексирание – съставяне на търсещ образ на документа или на предписание за търсене. Те са линейни отношения между лексикалните единици на информационния език, установени непосредствено при свързването на думи и/или словосъчетания във фрази. Чрез синтагматичните отношения се образува израз от високо ниво на интеграция, извършва се преход от дума към словосъчетание, от проста предметна рубрика към сложна, от прост класификационен индекс към сложен или съставен. Например индексът от основната таблица на УДК 661.25 Производство на сярна киселина е свързан с понятие от помощните таблици .011 Изчисляване и образува нова единица, която изразява понятието 661.25.011 Изчисляване на процесите при производство на сярна киселина.

Ако парадигматичните отношения са смислови връзки между понятията и имат постоянен характер, то синтагматичните са ситуативни, възникват в момента на отразяване на съдържанието на конкретен документ или информационното запитване.

Синтагматичните отношения стесняват обхвата на търсенето чрез поясняване и увеличават точността.

Свързването на лексикалните единици и получаването на понятие с по-високо ниво на интеграция става чрез прилагането на специални синтактични правила, нар. граматика на информационния език.

Граматика на информационния език.

Граматиката на информационния език представлява набор от правила, методи и средства за създаване на търсец образ на документ или за формулиране на предписание за търсене. В един информационен език могат да се използват различни граматични средства, изследвани и представени обобщено от дългогодишния изследовател на информационните езици Н. Гендина²¹⁰:

а) „Свободна граматика” всъщност означава минимално използване на правила. Свежда се до изброяване на лексикалните единици, включени в търсецния образ на документа или в предписанието за търсене без да има значение техния ред – само наличието им е достатъчно да покаже текстуална връзка между тях. Типичен пример за език със свободна граматика е езикът на ключовите думи. Индексирането на документите се реализира чрез причисляване на думи или словосъчетания от текста на документа (може и от заглавието и/или резюмето), които дават представа за съдържанието и служат за търсене, например: „техника на безопасност, авария, травматизъм”.

Свободната граматика като допълващо граматично средство може да се използва и в езици със сложна граматика, например при индексиране на съдържание на документ или запитване с УДК, се използва знакът плюс „+” за свързване на две и повече понятия, което предполага присъствието на лексикалните единици в един търсец образ, без да се посочват техните смислови отношения.

б) „Позиционна граматика” – предполага установяване на строг ред на представяне на лексикалните единици в търсецния образ на документ и/или предписанието за информационно търсене. В езиците от класификационен тип, например УДК, позиционната граматика се проявява в правилата за изграждане на сложен класификационен индекс, според който първо се поставя индексът от основните таблици, а след това определител/и от помощните таблици.

Позиционната граматика може да се реализира чрез „анкетен метод” за съставяне на търсец образ на документи. Той предполага използването на фиксирана схема - анкета за

²¹⁰ Гендина, Н. И. Цит. съч., с. 61–66.

подреждане на лексикалните единици, която определя реда на тяхното записване в търсения образ. Този метод на индексване е приложим в масив с еднотипна структура на документите. Например за унифицирано индексване на научно-технически документи обикновено се използва следната схема на анализ²¹¹:

1. Предмет или тема на изследване.
2. Област на приложение или използване.
3. Метод на изследване.
4. Условия, в които е протекло изследването.
5. Технически средства или оборудване, използвани в изследването.
6. Резултати от изследването.

Предимствата на този метод са компактност на записа, удобство да се съпоставят изразите при търсене, намаляване на субективността при индексване. Въпреки това, предвид голямото разнообразие от информация, съдържаща се в текстовете, този метод губи предимствата си, когато много от позициите в анкетата останат непопълнени при индексване.

в) „Показатели на връзката” е граматически метод, при който се включват символи (букви или цифри) в търсения образ на документа или информационното запитване, за да се покаже нагледно съществуващата логическа връзка между лексикалните единици. Този метод е резултат от развитието на първите посткоординатни езици с лексика, която се състои само от отделни думи (езикът „УниTERM”). Използването на това граматично средство се налага, поради обстоятелството, че отделните думи могат да образуват не едно, а повече смислови конструкции. Например, ако съдържанието на документ се изразява с ключовите думи: производство, нефт, рафиниране, автомобилни горива, могат да възникнат различни смислови парадигми – производство на нефт и рафиниране на автомобилни горива, но също и производство на автомобилни горива и рафиниране на нефт. За постигане на точност на търсенето се налага лексикалните единици да бъдат свързани посредством показатели, които указват връзката между тях – производство (А), автомобилни горива (А), рафиниране (Б), нефт (Б).

г) „Показатели на роля” е граматично средство, чрез което се посочва смисловата функция на отделната лексикална единица в търсения образ на документа или запитването. Те се посочват в определен ред. Например в посткоординатен език могат да се използват следните показатели на роля:

²¹¹ Пак там, с. 63.

P1 Обект на разглеждане.

P2 Област, сфера на приложение.

P3 Метод.

P4 Част, елемент, машина, съоръжения, вещества и т.н.

„Показателите на роля” ограничават обема на понятието, обозначено с лексикалната единица и повишават точността на търсенето. Въпреки че този метод е създаден и се използва основно в посткоординатните езици, все пак се прилагат сходни граматични средства и в езиците от класификационен тип. Например в УДК функцията на показатели на роля изпълняват общите (за място, за време, за процеси, за лица и т.н.) и специалните определители. В азбучно-предметната класификация указатели на роля са подрубриките (тематични, географски, хронологични и формални подрубрики). Те ограничават обема на понятието, изразено в предметната рубрика и съответно, увеличават точността. За сравнение предлагаме предметните рубрики 1. Училища и 2. Училища – България – Възраждане.

Методите „показатели на роля” и „показатели на връзка” се прилагат с цел подобряване степента на точност на информационното търсене. Според изследвания показателите за връзка увеличават точността на извличането на информация с до 10%, а съвместното използване на двата метода с 15 до 17%²¹².

Макар че езиците, които използват тези методи позволяват описване съдържанието на документите и/или запитванията с голяма точност, те са свързани и с някои недостатъци. Основният от тях е, че са изключително трудни за работа, изискват сложни процеси за индексирание и търсене.

д) „Грамматика на логическите оператори” не представлява граматически правила за индексирание, но се използва при търсене. Грамматиката на логическите оператори се основава на използването на логическите връзки И, ИЛИ НЕ, които позволяват логическо умножаване, събиране или изваждане на лексикални единици, т.е. изпълнението на операции от алгебра на логиката (булева алгебра). Логическите оператори служат за ограничаване обхвата на търсене. Те се използват само при съставянето на предписания за търсене и определят стратегията за търсене. Чрез тях могат да се съставят разнообразни логически изрази за точно формулиране на дадено информационно запитване²¹³.

²¹² Гендина, Н. И. Цит. съч., с. 65.

²¹³ Койчева, Елена. Информационна евристика: систематичен план за провеждане на информационно търсене в уеб ресурси. В: *Годишник на Софийския университет „Св. Климент Охридски”. Философски факултет. Библиотечно-информационни науки*, 8, 2016, с. 297. Достъпно онлайн на: <https://research.uni-sofia.bg/handle/123456789/1404>

След като разгледахме структурните компоненти на информационните езици следващата стъпка е да представим технологията на тяхното съставяне, т.е. изграждането на лексикалната и семантична основа на информационните езици. Поради ограничения в обема на това изследване ще представим общ универсален шаблон за създаване на информационен език. Той ще бъде полезен за разработчиците на лингвистично осигуряване на електронна библиотека, при необходимост от създаване на информационен език за решаване на задачи в конкретна електронна библиотека.

3.2.2. Технология за създаване на лексикално-семантична основа на информационен език

Конструирането на информационен език или създаването на неговата лексикално-семантична база включва следните основни етапи:

1. Събиране на понятия.
2. Лексикографска и семантична обработка.
3. Систематизиране на лексикалните единици.
4. Структурно оформяне на речниковия състав на езика.

Първи етап. Разработването на информационния език започва със събирането на специалните понятия от съответната тематична област на електронната библиотека. Това е най-важният етап в създаването на всеки информационен език, тъй като от качествено изпълнение на този етап до голяма степен зависи семантична сила на езика. Този етап задължително се предхожда от проучване на системата от знания и анализ на терминологията в съответната научна дисциплина или област. Затова в екипа на съставителите на информационен език се включват не само информационни специалисти, библиотекари, библиографи, лингвисти, но и експерти в предметната област.

В лексиката на информационния език се включват думи и/или словосъчетания, които носят смислова натовареност. Пренебрегват се служебни думи (частици, предлози, съюзи, междуметия), често използвани общонаучни термини, термини, които не са свързани с понятийния апарат на предметната област, за която се създава език. Подборът се основава на термини, които могат да бъдат представени в три категории: а) специализирани термини, които се употребяват само в определената област на знанието; б) термини от сродни или близки области на знанието; и в) общонаучни термини. Причисляването на всеки конкретен термин към речниковия състав на езика трябва да се основава на прагматични съображения като тематичния обхват на електронната библиотека и потенциалните възможности за възникване на запитване. Оптималният състав на речника е от значение за качеството на

индексирането и търсенето. Пропускането на понятия и невключването им като лексикални единици води до неизбежна загуба на информация, а прекомерното разширяване на речниковия състав прави езика труден и неудобен при търсене.

В практиката са наложени два основни метода за събиране на понятия за съставяне на речниковия състав на информационния език: дедуктивен и индуктивен²¹⁴.

Дедуктивният метод предполага използването на терминологични ресурси (справочници, стандарти и др.п.) и нормативни речници на съществуващи вече информационни езици. Основното преимущество на този метод е бързото натрупване на масив от специални понятия, в по-голяма или по-малка степен в нормирана форма и с изяснени семантични връзки. Така събраните понятия се отличават с относителна стабилност (наложени във времето). Като недостатък на дедуктивния метод се сочи известния риск в речниковия състав да не попаднат най-нови термини, ако не се използват най-актуални справочници.

Индуктивният метод се състои в свободното индексиране на първични или вторични документи. Събирането на понятията по метода на свободното индексиране позволява в речника да се включат нови и по-специфични термини. Предимствата на този метод, е че може да бъде реализиран и автоматично. Освен това позволява да се получат данни за честотата на използване на думи, което подпомага избора на предпочетен термин при синонимия.

За да се използват предимствата на двата метода при реалното разработване на информационен език, те могат да се приложат съвместно. Така като основни източници за подбор на лексика на информационния език се включват:

1. Съществуващи информационни езици, които разполагат с контролиран речник: класификационни таблици на УДК, на ДКД, на ББК, списъци на предметни рубрики, класификатори, тезауруси и др.
2. Терминологични и тълковни речници, тематични справочници и енциклопедии.
3. Стандарти и други нормативни издания.
4. Първични и вторични документи по тематиката, за която се създава информационният език, подлежащи на свободно индексиране. Обемът на документите се препоръчва²¹⁵ да бъде 3–4 хиляди. Такъв обем дава възможност да се извлече представителен брой понятия, които в достатъчна степен характеризират предметната

²¹⁴ Габровска, Свободозаря, Пейчева, Веселина. Методични и организационни проблеми при съставянето на тезауруси в обществените науки. В: *Съвременни проблеми на научната информация*. София: БАН, 1974, с. 34–35.

²¹⁵ Пак там.

област. Хронологичния обхват на документите трябва да е от последните 5–6 години, за който период се актуализира терминологията.

5. Набор от запитвания на потребители, които обикновено съдържат най-новите термини.

На този първи етап от разработването на информационен език е много важно да се постигне оптимална пълнота в обхвата на речниковия състав.

Вторият етап от разработването на информационен език предполага обработката на събраните понятия. Обработката на лексиката обхваща два различни процеса: лексикографска обработка и семантична обработка.

Лексикографската обработка има две основни задачи: а) установяване на унифицирана форма за записване на лексикалните единици; б) решаване на въпроса относно формулирането на лексикалната единица като отделна дума (унитерм) или като словосъчетание.

Първата задача обхваща дейности свързани с нормализацията и правописа на думите като: последователно прилагане на правилата за единствено или множествено число, за използване на съкращения и др. Правила, свързани с българския език, наложени в практическия опит по създаването на конкретни информационни езици могат да се намерят в някои източници²¹⁶. Правописът на включените термини следва да отговаря на възприетите в страната правописни норми. За нашата страна те са представени в „Официален правописен речник на българския език”²¹⁷.

Една от основните трудности в целия процес на изграждане на речниковия състав на информационен език е решаването на въпроса, кога дадено словосъчетание трябва да влезе като цялостна единица и кога следва да бъде разложено на съставлящите го думи. От една страна, въвеждането на словосъчетания води до увеличаване обема на речника, от друга, разлагането на всички словосъчетания до думи може да доведе до загуба на специфичното им съдържание, (например: база от данни, висок печат, първа помощ). Вредата от него се проявява особено при информационното търсене, поради възможността от образуване на грешни комбинации на съставлящите елементи, което значително понижава точността на резултатите от търсенето. Следователно трябва да се търси оптимално решение и да се изработят правила, които да определят начина на въвеждане на словосъчетанията в речника

²¹⁶ Габровска, Свободозаря, Пейчева, Веселина. Цит. съч.; Атанасова, В. Цит. съч.

²¹⁷ Българска академия на науките. *Официален правописен речник на българския език*. София: Просвета, 2012, 676 с.

на информационния език. Опитът показва, че могат да се използват следните ограничения²¹⁸:

а) Словосъчетанието се използва като единна лексикална единица, когато:

– представлява общоприет термин и разлагането му на съставляващите компоненти води до загуба на специфичността му, напр. Организационно поведение, Твърди тела, Висок печат и др.;

– отделните съставляващи компоненти се използват, толкова често самостоятелно, че при информационно търсене би се получил голям резултат от попадения, например База от данни;

– никой от съставляващите компоненти не се съдържа в речниковия състав на езика като самостоятелно понятие, например Първа помощ, Висок печат.

б) Словосъчетанието се заменя с образуващите го единични думи, когато логическото произведение на съставляващите понятия точно предава неговото специфично съдържание и словосъчетанието не се използва често при индексирание и търсене, например Строеж на материята – строеж, материя.

В случай, че е решено в разработвания информационен език да се използват като лексикални единици не само единични думи, но и словосъчетания, е необходимо да се определи реда на записване на думите в словосъчетанията. По отношение на словосъчетанията, които се състоят от прилагателно и съществително, се изисква решение – в прав ред (прилагателно и съществително) или в инверсия те да бъдат използвани като лексикални единици. Инверсията позволява да се групират близки по съдържание понятия и да се формират по-широки тематични комплекси.

Лексикалните единици на този етап от разработването на информационния език се подлагат и на семантична обработка. За достигането на висока точност и пълнота на информационното търсене е необходимо въвеждането на контрол върху речниковия състав на информационния език. Той представлява съвкупност от методи и средства за отстраняване на синонимията, омонимията и полисемията и фиксиране на парадигматичните отношения между лексикалните единици.

Отстраняването на синонимията е много важно условие за постигане на пълнота на търсенето. Това става чрез определянето на приет термин за индексирание и означаване на връзката му със синонимни термини.

²¹⁸ Габровска, Свободозаря, Пейчева, Веселина. Цит. съч., с. 39.

Елиминирането на многозначността (омонимията и полисемията) при стандартизирането на лексиката се постига по два начина. Първият, най-простият начин, е да се разгърне многозначната дума до словосъчетание, например Мрежи – Компютърни мрежи, Риболовни мрежи. Вторият метод предполага използването на т. нар. определители към многозначната дума, например: Време (астрономия), Време (метрология), Време (физика), Време (философия).

На *третия етап* от разработването на информационен език се извършва систематизиране на лексикалните единици. То се основава на парадигматичните връзки между тях. Следователно най-важната задача, по време на етапа на систематизиране на лексиката, е да се идентифицират видовете парадигматични отношения между избраните термини.

От множеството разнообразни парадигматични отношения следва да се определят тези с най-голямо значение за информационното търсене:

а) Отношения на еквивалентност (синонимия) -, съвпадение на думите по смисъл, (например пълно название и означение с абrevиатура Информационни и комуникационни технологии и ИКТ), лексикална синонимия (две думи със сходно значение и с различни корени: Безредие и Анархия), логическа синонимия (термин и определение, например Столица – Най-главният град на една държава).

б) Йерархични отношения – отношения на подчинение, когато едно понятие подчинява или включва други понятия: „род-вид” и „цяло-част”.

в) Логически отношения и асоциации – отношения, при които едно понятие е смислово свързано с друго, например предмет – свойство (Метал и Електропроводимост), предмет – процес (Метал – Леене), предмет – предназначение (Филтър – Филтриране на примеси), причина – следствие (Авария – Злополука) и др.

В резултат на идентифицирането на парадигматичните отношения става възможно да се изградят групи от понятия, свързани логически. От тях се избира една смислова доминанта, която представлява цялата група от гледна точка на информационното търсене. Например подход прилаган при съставянето на схема на понятията в език от дескрипторен тип предлага следния алгоритъм: 1) определяне на основното значение на дескриптора; 2) включване на синонимите на дескриптора в дескрипторната група; 3) последователно включване на родовете, видовете и асоциативните дескриптори. В резултат на прилагането на алгоритъма се образува дескрипторна група.

Този подход за систематизация на понятията е приложен в тезауруса, използван в електронния каталог на НБКМ*. По-долу е представен пример на дескрипторна група от този тезаурус:

ВОДОПРОВОДИ [*групов дескриптор*]

ВЖ ОТ Водопроводни инсталации [*синонимна ключова дума, недескриптор*]

ВЖ ОТ Водопроводни мрежи [*синонимна ключова дума, недескриптор*]

ВЖ И ОТ Водоснабдяване [*родов дескриптор*]

ВЖ И ОТ Канали [*родов дескриптор*]

ВЖ И Водопроводни тръби [*видов дескриптор*]

ВЖ И Напорни тръби [*видов дескриптор*]

ВЖ И Помпени станции [*видов дескриптор*]

ВЖ И Помпи [*видов дескриптор*]

ВЖ И Рилски водопровод [*видов дескриптор*]

ВЖ+ВЖ И ОТ Тръбопроводи [*асоциативен дескриптор*]

Подходът позволява да се получи строго подредена понятийна система, да се визуализират семантичните отношения, обединяващи лексикалните единици и да се използват в процеса на търсене. Означаването на родово-видовите отношения е особено важно при т. нар. допълнително индексирание. С цел да се увеличи пълнотата и точността на информационното търсене, търсещият образ на документите или предписанията за търсене се допълват с понятия разположени в йерархичните схеми на по-високо (възходящо допълнително индексирание) или на по-ниско равнище (низходящо допълнително индексирание). Обикновено възходящо индексирание се прилага при индексирание на документи, а низходящото – на информационните запитвания. Указването на асоциативните връзки дава възможност за увеличаване на вариантите на предписанията за търсене и има основно значение при определяне стратегията на информационното търсене.

Последният, *четвърти етап* е организационното оформление на лексиката и графичното представяне, което зависи изцяло от типа на информационния език. В зависимост от езика, вече нормализираната и систематизираната лексика ще бъде трансформирана в предметни рубрики, в дескриптори, или в класификационни индекси и т.н., представени по определен начин. Например лексико-семантичната основа на езика може да бъде представена във вид на йерархична класификационна таблица (УДК, ДКД,

* <http://www.bg.cobiss.net/>

ББК и др.), на рубрикатор (напр. Государственный рубрикатор научно-технической информации, ГРНТИ), на класификатор, на списък на предметни рубрики, на тезаурус и др. Важно е да се отбележи, че структурното оформление на лексиката, според типа на информационния език трябва да бъде унифицирано.

3.2.3. Типологизация на информационните езици

Представянето на информационните езици в типологизация има за цел да очертае най-съществените характеристики и отличителни особености на видовете информационни езици, да разкрие предимствата и недостатъците при тяхното практическо приложение, и да подпомогне изборът на необходими информационни езици за лингвистично осигуряване на електронна библиотека.

Разглежданата тук типологизация е често коментирана и наложена в теоретичната литература още от по-ранните изследователи на информационните езици. Типологизацията се основава на най-съществените признаци на лексиката на информационните езици: начин на представяне на лексикалните единици, етап на координация на лексикалните единици, начин на задаване на лексикалните единици.

а) Информационни езици, според начина на представяне на лексикалните единици

Най-характерния признак, по който могат да се типологизират информационните езици е начинът на представяне на лексикалните единици. Възможни са два способа: при първия се използват специални кодове (цифрови или буквено-цифрови), а при втория – думи и словосъчетания, взети от някой от естествените езици. На основание на този признак се различават следните видове информационни езици: кодови информационни езици и словесни информационни езици.

В *кодовите информационни езици* лексикалният състав е представен чрез кодове (индекси). Следователно изразът на лексикалната единица не съвпада със значението, например: 51 Геометрия (УДК) или 47.13.07 Технология и оборудване за производство на прибори и устройства за наноелектрониката (ГРНТИ). Затова кодовите езици задължително разполагат със специални двузначни „речници”, които позволяват кодовите нотации да се „превеждат” на естествен език и обратно. Типичен пример за такива речници са класификационните таблици с азбучно-предметния им показалец, които осигуряват преход от естествен език на кодов или обратно. На този принцип са създадени УДК, ДКД, ББК, МПК, МКС и други класификатори и рубрикатори.

Словесните информационни езици използват за представянето на лексикалните единици думи и словосъчетания от естествен език. Към словесните информационни езици се числят езикът на предметните рубрики, езикът на ключовите думи, дескрипторният език, езикът на библиографското описание.

б) Информационни езици, според етапа на координацията на лексикални единици

Лексикалната единица в някои информационни езици може да бъде не само отделна дума или устойчиво словосъчетание, но цяла фраза, дори изречение. Примери: 640.4 Управление и организация на хотели и заведения за обществено хранене (УДК); А44С27/00 Изготвяне на ювелирни изделия (МПК); Децата като художници (Тезаурус НБКМ). Според това на какъв етап се извършва свързването на понятия, за да изразят по-специфични такива, информационните езици се делят на предкоординатни и посткоординатни.

Предкоординатни езици са езиците, чийто речников състав се състои от имена на сложни класове, образувани преди индексирването или по време на индексирването²¹⁹. Отличителната особеност на езиците от предкоординатен тип се състои в това, че при тях предварително са фиксирани и представени имената на основни класове и подкласове, изразени не само с помощта на отделни думи и словосъчетания, а и с разгърнати словосъчетания и цели фрази. В тези словосъчетания и фрази думите се свързват (координират) една с друга преди индексирването, т.е. те са предварително координирани (предкоординатни). Предкоординатни езици са например езика на предметните рубрики и кодовите класификационни езици: УДК, ДКД, ББК, МПК, МКС, класификатори, рубрикатори и др. В основата на предкоординатното индексирване е принципът за представяне на основното съдържание на документ или запитване като цяло и отразяване на предмета посредством лексикални единици от предварително съставен списък, от които трябва да се изберат най-подходящите по смисъл индекси (код или предметни рубрики).

При *посткоординатните (координатни) езици* търсещите образи на документите и предписанията се образуват от думи или устойчиви словосъчетания чрез логическо умножение по време на индексирването на документа и запитването при провеждането на информационното търсене²²⁰. Езиците от посткоординатен тип са езикът на ключовите думи и дескрипторният език. Те се състоят предимно от отделни думи и словосъчетания, които се свързват (координират) по време на индексирване.

²¹⁹ Гендина, Н. И. Цит. съч., с. 88.

²²⁰ Пак там.

Съществените характеристики на предкоординатните и посткоординатните езици са анализирани от дългогодишния изследовател и експерт в различни работни групи на ИФЛА Елайн Свенионий²²¹.

в) Информационни езици, според начина на избор на лексикалните единици

В зависимост от това как се избират лексикалните единици за индексирание на документ или запитване, информационните езици се делят на езици с контролирана лексика и езици с неконтролирана лексика.

В *информационните езици с контролирана лексика* всички лексикални единици се представят в специален речник (нар. нормативен речник). Нормативният речник съдържа възможните термини за индексирание и определя синтактични правила за тяхното съчетаване²²². Такива речници представляват, например таблиците на УДК, ДКД, ББК, рубризаторите, тезаурусите, списъците на предметни рубрики. Присвояването на термини за индексирание се „контролира или изпълнява от човек“²²³.

Речникът на езиците с контролирана лексика не само поддържа процеса на индексирание на съдържанието на документите, но и търсенето като: повишава вероятността индексаторът и потребителят да отразяват едно и също понятие еднакво; индексаторът и потребителят да могат да стигнат до търсеното понятие (тема, предмет) чрез системата от връзки – (родови термини, видови термини, асоциативни термини); осигурява последователно индексирание; служи на потребителя като справочник, чрез който да се ориентира в предметната област; формира основа за организация на знанията в класификационна схема за удобно навигиране²²⁴.

Информационните езици с контролирана лексика са задължително средство в лингвистичното осигуряване на електронна библиотека, тъй като подобряват ефективността на търсенето и осигуряват ценна интелектуална помощ на потребителите при търсене. Те осигуряват намирането на документи, дори потребителя да не е запознат с темата²²⁵.

²²¹ виж например: **Svenonius**, Elaine. *The Intellectual Foundation of Information Organization*. Cambridge, Mass.: MIT Press, 2000.

²²² **Broughton**, Vanda. *Essential thesaurus construction*. London: Facet Publishing, 2006.

²²³ **Rowley**, Jennifer. The controlled versus natural indexing languages debate revisited: A perspective on information and retrieval practice and research. In: *Journal of Information Science*, 1994, Vol. 20, № 2, p. 108–119.

²²⁴ **Olson**, Hope A., **Boll**, John J. *Subject analysis in online catalogs* [online]. Englewood, Colorado: Libraries Unlimited, 2001. [Viewed 20.06.2018]. Available from: https://archive.org/details/subjectanalysisi00alu_34g

²²⁵ **Mann**, Thomas. Will Google's keyword searching eliminate the need for LC cataloguing and classification?. In: *Internet Archive* [online]. Last updated on August 15, 2005. [Viewed 20.06.2018]. Available from: <http://www.guild2910.org/searching.htm>

Въпреки безспорните предимства на езиците с контролирана лексика, те имат и някои недостатъци. Като основни ограничения се сочат високата стойност и трудоемкост на индексирането и поддържането на речник²²⁶.

С неконтролирана лексика е езикът на ключовите думи. Той използва естествен език, чието приложение може да се формализира чрез специални правила (род, число и др.п.), но не се съставя нормативен речник на лексикалните единици. Описанието на документа става непосредствено чрез думи, използвани в текста и/или заглавието, и/или резюмето на документа.

Представената типологизация разделя информационните езици на типове и видове с оглед анализа на лексиката на езика по различни критерии. Тя е общоприета и наложена в теорията на информационното търсене, но за целите на настоящото изследване е неудобна. Затова предлагаме друг подход на разглеждане на информационните езици, с оглед тяхната роля и специфика в лингвистичното осигуряване на електронна библиотека.

3.3. Предназначение и специфики на информационните езици в лингвистичното осигуряване на електронната библиотека

За постигане на крайната цел на настоящето изследване обоснован избор на информационни езици за конструиране на модел на лингвистично осигуряване на българска електронна библиотека с универсална тематика, намираме за удачно езиците да бъдат разгледани в 3 основни групи. Тези групи обединяват езици, според това по какъв начин различните информационни езици представят и описват документите: цялостно представяне на документите, представяне на предмета (темата) на документите, многоаспектно представяне на съдържанието на документите.

3.3.1. Информационни езици за цялостно представяне на документите

Цялостното отразяване на документите в електронна библиотека се извършва посредством набор от данни, характеризиращи документа по определени признаци, които е прието да бъдат наричани метаданни²²⁷. Метаданните са централен компонент на лингвистичното осигуряване на всяка информационна система (в частност електронна

²²⁶ **ИФЛА.** Руководство по обеспечению тематического доступа в национальной библиографии. В: *IFLA* [онлайн]. Санкт-Петербург, 2017. [Прегледан 20.06.2018] Достъпно от: <https://www.ifla.org/files/assets/hq/publications/series/45-ru.pdf>

²²⁷ **Когаловский,** М. Р. Метаданные в компьютерных системах. В: *Программирование*, 2013, № 4, с. 29. Достъпно онлайн на: <http://www.ipr-ras.ru/articles/kogalov13-03.pdf>

библиотека). Тяхната роля се сравнява с ролята на библиотечния каталог за управлението и организирането на библиотечния фонд. Съответно метаданните реализират основни функционалности на електронната библиотека:

- въвеждане, обработка, организация и съхранение на информационните обекти;
- достъп до информационните обекти;
- търсене на отделен информационен обект или търсене на съвкупност от документи, обединени от общ признак;
- навигация в информационното пространство по тема, автор, заглавие и т.н.;
- управление на правата за достъп до информацията, включително защита на авторското право, организация на плащанията за достъп и др.

Терминът „метаданни“ се разбира и определя по различен начин от множеството професионални организации, които създават, съхраняват или използват различни информационни системи и ресурси²²⁸. Според изследователи днес са се наложили два основни подхода²²⁹. Първият произхожда от библиотечната общност и нейният опит в създаването и поддържането на документални информационни системи и електронни каталози. Вторият подход за разглеждане на „метаданни“ идва от компютърните науки и е свързан с развитието на базите от данни и други технологии в информатика за управлението на данни и знания.

Според М. Когаловский „метаданни“ в библиотеките и в технологията на текстовото търсене се свързва с понятия като: набор от значения, индексирани свойства на документите, библиографско описание на документи, както и с по-частните понятия – идентификатори на текстови документи, класификационни рубрики, ключови думи и др.²³⁰

В литературата най-често метаданните се разглеждат в три категории, с оглед тяхното предназначение: описателни, структурни и административни метаданни. Описателните метаданни характеризират общо документа. Пример за описателни метаданни са данните на библиографското описание или в частност множеството от елементи на Dublin Core. Структурните метаданни представят структурата на описвания ресурс (елементи на структурата, последователност на страници и др. п.).

²²⁸ **Миланова**, Милена. *Българската каталогизация в глобалното...*, с. 80.

²²⁹ **Burnett**, Kathleen, Ng, Kwong Bor, **Park**, Soyeon. A Comparison of the Two Traditions of Metadata Development. In: *Journal of the American Society for Information Science* [online], Vol. 50, 1999, Issue 13, p. 1209–1217. [Viewed 20.06. 2018]. Available from: https://www.researchgate.net/publication/220433138_A_Comparison_of_the_Two_Traditions_of_Metadata_Development

²³⁰ **Когаловский**, М. Р. Метаданные в компьютерных системах. Цит. съч., с. 28.

Административните метаданни включват сведения необходими за управлението на документите (място на съхранение, носител на права, условия за достъп и др.).

По-детайлна класификация на видовете метаданни и тяхното конкретно предназначение предлага Националната организация по стандартизация в САЩ (National Information Standards Organization, NISO) в публикация, предназначена да бъде ръководство за създатели и потребители на метаданни²³¹ (виж табл. 6).

Таблица 6. Типове метаданни.

Тип метаданни		Примерни свойства	Предназначение
Описателни метаданни		Автор, Заглавие, Дата на публикуване, Вид на документа (жанр), Предмет, Ключови думи и др.	Идентифициране, Търсене, Навигиране, Представяне, Семантична съвместимост
Технически метаданни		Тип на файла, Размер на файла, Дата на създаване и др.	Управление на информационните обекти, Съхранение, Използване
Метаданни за съхранение		Място на съхранение, Условия за използване (разплащане) и др.	Управление на информационните обекти, Съхранение, Използване
Метаданни за права		Статус на авторските права, Лицензни условия, Носител на права и др.	Управление на информационните обекти
Структурни метаданни		Структура на документа и др. п.	Преглеждане на документа

За целите и задачите на настоящото изследване ние ще се съсредоточим върху описателните метаданни, защото именно те служат за представяне на характеристиките на документите в електронна библиотека с цел търсене и идентифициране. Обаче, като вземаме под внимание широкото използване на понятието „метаданни”, както и разнообразието от типове и видове метаданни, смятаме, че е по-подходящо да използваме термина библиографско описание. Библиографското описание се определя като набор от

²³¹ NISO. *Understanding metadata: What is metadata and what is it for?* [online]. National Information Standards Organization (NISO), 2017, p. 7. [Viewed 20.06.2018]. Available from: https://groups.niso.org/apps/group_public/download.php/17446/Understanding%20Metadata.pdf

структурирани данни, с които се описват характеристиките на информационни обекти с цел тяхното идентифициране, откриване и управление. Данните могат да бъдат кодирани и формализирани с цел машинна обработка*. Необходимите данни за съставяне на библиографско описание обикновено се взимат от специфични източници в документа.

Библиографското описание трябва се съставя в съответствие с международно приети стандарти и правила²³². Обикновено те регламентират елементите от данни, които трябва да се запишат, значенията на елементите от данни, последователността на представяне на елементите и др. Общоприетите стандарти и правила имат за цел да осигурят унифицирани и общоразбираеми описания, да опростят обмена на информация и в крайна сметка да направят възможна интеграцията на описания от различни източници в единна информационна система. Примери за такива стандарти са: „Международният стандарт за библиографско описание” (International Standard Bibliographic Descriptions, ISBD), „Международен стандарт за архивни описания” (International Standard Archival Description, ISAD), „Описание на ресурс и достъп” (Resource Description and Access, RDA) и др.

Специално за библиотечната общност библиографското описание се регламентира от „Международният стандарт за библиографско описание”²³³. Той определя изискванията за описание и идентификация на най-често срещаните публикувани документи, които могат да влязат в библиотечни колекции, включително електронни информационни обекти²³⁴. „Международният стандарт за библиографско описание. Консолидирано издание” е приложим за библиотечни системи в онлайн режим²³⁵. В таблица 7 е представена структурата и елементите от данни на библиографско описание съгласно Международният стандарт за библиографско описание. Консолидирано издание²³⁶.

* Такова е значението на описателните метаданни - структурирани и кодирани данни, които описват характеристиките на информационните обекти, с цел тяхното идентифициране, откриване и управление виж: **Committee** on Cataloging: Description and Access. Task Force on Metadata. Final Report [online] CC:DA/TF/Metadata/5 June 16, 2000. In: *ALCTS: Association for Library Collections & Technical Services*. [Viewed 21.06.2018]. Available from: <http://downloads.alcts.ala.org/ccda/TF-meta6.html>

²³² ИФЛА. Изложение на международните принципи на каталогизацията. 2009, с. 4.

²³³ Пак там.

https://www.ifla.org/files/assets/cataloguing/icp/icp_2009-bg.pdf

²³⁴ ИФЛА. *ISBD: Международният стандарт за библиографско описание. Консолидирано издание*. София: Българска библиотечно-информационна асоциация, 2012, с. 1. Достъпно и от:

<https://www.ifla.org/files/assets/hq/publications/series/44-bg.pdf>

²³⁵ Пак там, с. VIII.

²³⁶ Пак там, с. 11–14.

Таблица 7. Структура и елементи от данни на библиографското описание.

Области на ISBD	Елементи, последователности и пунктуация
Област 0. Област на формата на съдържанието и вида на средството	Форма на съдържанието () Определение на съдържанието ; Следващо определение на съдържанието : Вид на средството
Област 1. Област на заглавието и сведенията за отговорност	Основно заглавие = Паралелно заглавие : Допълнение към заглавието / Сведение за отговорност ; Следващо сведение ; Следващо заглавие от същия автор . Следващо заглавие от друг автор и т.н.
Област 2. Област на изданието	Сведение за изданието / Сведение за отговорност ; Следващо сведение
Област 3. Област на специфична характеристика на материала или на вида на ресурса	Математически данни Сведение за мащаб ; Сведение за проекция () Сведение за координати ; Сведение за равенство Сведение за музикален формат Номерация Цифрово означение () Хронологично означение = система за алтернативно означение ; нова поредица на номерацията
Област 4. Област на публикуването, произвеждането и/или разпространяването и др.	Място на публикуване ; Следващо място : Име на издателя, производителя и/или разпространителя , Дата на публикуване и/или произвеждане () Информация за отпечатването или произвеждането Място на отпечатване или произвеждане : Име на печатаря или производителя , Дата на отпечатване или произвеждане
Област 5. Област на описание на материала	Обем : Други физически подробности ; Размери + Сведение за съпроводителен материал
Област 6. Област на серията и на многочастния монографичен ресурс	Основно заглавие на серия : Допълнение към заглавието на серия / Сведения за отговорност, отнасящи се до серия ; Следващо сведение , Международен стандартен номер на серия ; Номерация в серия
Област 7. Област на забележките	няма елементи

Област 8. Област на идентификатора на ресурса и условията за разпространяване	Идентификатор на ресурса = Ключово заглавие : Условия за разпространяване
---	---

За създаване на бази от данни и онлайн библиотечни каталози се използва специален формат за записване на данните за документите – формат MARC (от MACHine READable Cataloging). Най-разпространените MARC формати - MARC21 и UNIMARC са съобразени със стандартите за обмен на информация: *ISO 2709 Информация и документация. Формат за обмен на информация** и *ANSI/NISO Z39.2 Information Interchange Format*.

Друг пример за международен стандарт за описание на ресурси с широко приложение в уеб е ISO 15836-1:2017 Множество от елементи метаданни Dublin Core.

Библиографското описание дава стандартен и унифициран набор от елементи от данни, по които може да се проведе търсене. Ето защо библиографското описание се разглежда от изследователите²³⁷ като специфичен информационен език, който служи за общо представяне на документа, т.е. неговите физически характеристики като автор, заглавие, дата, формални характеристики и т.н.

Азбуката на езика на библиографските данни съвпада с азбуката на естествения език. Освен това тя включва арабски и римски цифри, пунктуационни знаци (точка, запетая, тире, двоеточие, точка и запетая, многоточие), квадратни и кръгли скоби, символи (наклонена черта, две наклонени черти), математически знаци (знак плюс и знак за равенство). Знаците в библиографското описание имат специфична роля, както е в ISBD, където те служат като средство за разпознаване и представяне на елементите от данни и за тяхното разбиране, независимо от езика на описанието. Обикновено библиографското описание се записва на езика на документа.

Лексика. Лексикалните единици на езика на библиографските данни са елементите на библиографското описание (автор, заглавие, дата на издаване и т.н.). Елемент на библиографското описание е най-малката логическа единица от данни използвана в библиографското описание. Елементите се изразяват чрез думи или словосъчетания от естествения език. Доколкото лексиката в естествения език се характеризира с изменчивост и вариативност е целесъобразно в библиографското описание някои от елементите от данни да бъдат подложени на нормативен контрол, което осигурява по-висока степен на ефективност на търсенето. За тази цел се прилагат правила за стандартно представяне и унифициране на елементи - имена на автори, наименования на организации, дати и т.н. Те

* ISO 2709:2008 Information and documentation – Format for information exchange.

²³⁷ Антополски, А. Б. Цит. съч., с. 62–65.

могат да бъдат записани в специални списъци на допустими формализирани стойности - кодове на езици, на страни и нормативни файлове.

Елементите от данни в библиографското описание могат да бъдат задължителни и незадължителни. Елементите, които задължително трябва да бъдат посочени са от съществено значение за идентифицирането, търсенето и откриването на документите. В традиционните библиотечни системи за търсене (фишов каталог) броят на елементите, по които се провежда търсене е много ограничен. В автоматизиран режим по всеки от елементите на библиографското описание може да бъде организирано търсене, стига това да бъде целесъобразно.

Парадигматичните отношения в езика на библиографските данни не са изразени явно. Самата структура на библиографското описание предопределя йерархични отношения. Те се изразяват във връзката между елементите и областите на библиографското описание или в MARC форматите между полетата от данни и елементите от данни. Област на библиографското описание (съответно поле) е по голяма структурна част на библиографското описание, в която са включени данни от определена категория или набор от категории.

Синтагматичните отношения в езика на библиографското описание се реализират под формата на позиционна граматика: строгата последователност в представянето на отделните елементи се регламентира от стандарта за описание.

Стандартизираната структура на библиографското описание е много важна за организирането на търсенето по отделни данни. Това е от съществено значение за машинночетимата форма на библиографско описание.

Конкретни примери на лингвистични средства, които служат за цялостно описание на документите в онлайн библиотечни системи са форматите за машинночетими библиографски записи – MARC и схеми за описание на ресурси като Dublin Core (DC), Metadata Object Description Schema (MODS) The Encoded Archival Description (EAD), Learning Object Metadata (LOM) и др.

3.3.2. Информационни езици за представяне на предмета на документите

Едно от изходните понятия, основополагащи за теорията и практиката на индексирването е понятието „предмет”. То представлява основната тема, която се разглежда

в документа. В теорията на информационните езици, както и в класическата логика предмет има значение на обект като материална вещ, явление, свойство или отношение²³⁸.

За представяне на предмета на документите и за тематичен достъп в библиотечна традиционна или електронна система се използват класификационните езици. Те са езици от предкоординатен тип и имат свойството да разделят обектите (понятията) в класове, описани чрез предварително свързани признаци. Те отразяват общо, комплексно съдържанието на документа и позволяват търсене по неговия предмет. Те са езици с контролирана лексика.

Класификационни езици могат да бъдат изградени като йерархични класификации и нейерархични класификации.

Йерархичните класификации организират понятията, включени в своя речник чрез отношения на йерархия и съподчинение и образуват единна класификационна схема, представена под формата на йерархично дърво. Йерархията е отношение на подчинение, при което един клас представлява подклас на друг, по-широк клас, т.е. йерархията е отношение между класове по вертикала. Съподчинението е отношение, което свързва класовете, които са подкласове на по-широк клас, т.е. това са отношения по хоризонтала между подкласовете на един и същ клас. Йерархичните класификации могат да бъдат реализирани в структури – като монойерархия или полийерархия. Монойерархия е йерархична структура, при която всеки клас има само един непосредствено подчинен клас. Полийерархия е йерархична структура, при която всеки клас има два или повече непосредствено подчинени класа.

Йерархичната класификация се основава на логическата операция деление на понятията. Признакът, по който се прави делението, се нарича основание за деление. Основният принцип за изграждане на йерархичната класификация е дедукцията – „от общото към частното”.

Структурата на йерархичната класификация се изгражда като изходното множество от понятия (нулево ниво) се дели според избран класификационен признак на класове, които образуват първото ниво. Всеки клас от първо ниво, според собствен класификационен признак, се разделя на подкласове, които образуват второто ниво. Класовете от второ ниво се делят съгласно свой признак за деление на подкласове, които

²³⁸ **Svenonius**, Elaine. Subject Languages: Introduction, Vocabulary Selection, and Classification. In: *The Intellectual Foundation of Information Organization* [online]. Cambridge, Massachusetts: MIT Press, 2000, p. 126–146. [Viewed 20.06.2018]. Available from: <https://sites.evergreen.edu/wp-content/uploads/sites/226/2016/08/svenonius-chap8.pdf>

образуват третото ниво и т.н. В резултат се образува разклонена система на много нива, обикновено представяна като дървовидна графика, блок-схема или таблица.

Цифровото или буквено-цифровото обозначение на всеки клас/подклас от йерархичната класификация, който има определено, място по отношение на другите класове, се нарича индекс или код. Той е придружен от наименование на понятието на естествен език.

Йерархичните класификации се характеризират с качества като дълбочина и обем. Дълбочината на класификацията е броят на класификационните нива според признаците, избрани в качеството на основание за деление. Дълбочината на класификационната система се определя от степента на конкретизиране на подкласовете. Съществуващите йерархични класификации имат различна дълбочина, например в УДК и ДКД тя е равна на десет, МКС и ГРНТИ имат три нива на деление. Обемът на йерархичната класификация е съвкупността от всички индекси, т.е. класовете и подкласовете на всички налични в дадената класификация йерархични нива. Обемът е един от най-важните показатели на семантичната сила на йерархичната класификация. Той характеризира богатството на речниковия състав, например в таблиците на УДК, само класовете са 70 000, общият брой индекси надхвърля 250 000²³⁹.

Често като важно качество на конкретни йерархични класификационни езици се сочи тяхната универсалност²⁴⁰. Това качество може да се отнася до: тематичния обхват и означава, че класификацията обхваща целия универсум от знания; до вида на документите; до областта на приложение (видовете информационни системи, в които може да се използва). В табл. 8 са представени някои от съществуващите йерархични класификации, според нивото им на универсалност.

Таблица 8. Йерархични класификации според ниво на универсалност.

Ниво на универсалност	Йерархични класификации
Универсални по тематика и видове документи	УДК, ДКД, ББК и др.
Универсалност по тематика (политематичност), но специализирани по вид на документите	МПК, МКС и др.
Универсални по тематика и видове на документите, но ограничени по сфера на приложение	Класификатор на архивните фондове в Националния архивен фонд; Книгоиздателски и книготърговски класификации ВIC's Subject Categories and Qualifiers Scheme, Thema; Класификации в интернет каталози (напр. Yahoo и др.).

²³⁹ UDC Master Reference File (MRF). In: *UDC Consortium* [online]. [Viewed 15.06.2018]. Available from: <http://www.udcc.org/index.php/site/page?view=mrf>

²⁴⁰ Антопольский, А. Б., Цит. съч., с. 98.

Йерархичните класификации се прилагат в библиотечно-информационни системи за тематично търсене на информация, за систематизиране и структуриране на информационните масиви.

Структура на йерархичните класификации като информационен език.

Азбуката на йерархичните класификации може да се състои от цифри или пък да бъде смесена от цифри и букви. Освен това азбуката може да съдържа пунктуационни знаци (точка, дефис, апостроф, двоеточие, запетая, наклонена черта, скоби), специални знаци и символи (стрелка, квадрат). Към йерархичните класификационни езици с цифрова азбука се числят УДК, ДКД, рубрикатори и класификатори и др. Със смесена, буквено-цифрена азбука са например МПК и ББК.

Лексиката във всички йерархични класификационни езици има унифицирана форма – състои се от комбинация от цифров или буквено-цифров код и словесен израз (нотация). В класификациите УДК, ББК, МПК лексикална единица е класификационният индекс и съответният му словесен израз – класификационна рубрика, в рубрикаторите и класификаторите лексикалната единица се нарича „код“ и също е придружена от рубрика. Въпреки разликата в терминологията, лексикалните единици в йерархичните класификации имат една и съща структура (табл. 9).

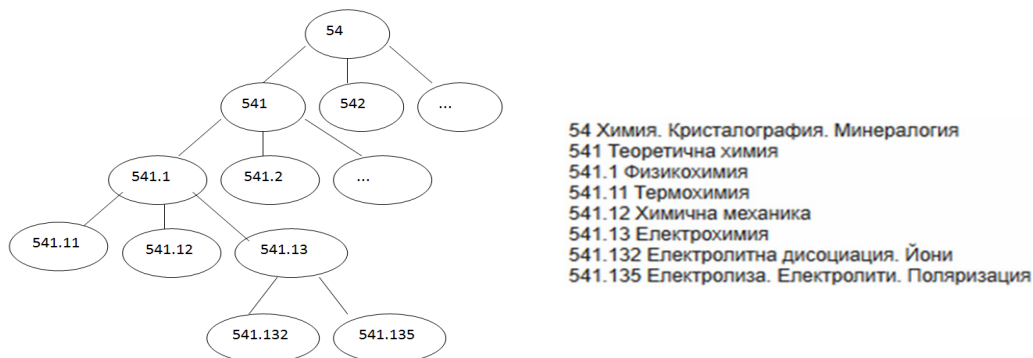
Таблица 9. Лексикални единици в йерархични класификации.

Лексикална единица			Информационен език
Код, индекс (израз)	индекс	Словесен израз (значение)	
701		Философия на изобразителното и декоративното изкуство	ДКД
004.738.5		Интернет. Глобална мрежа (World Wide Web)	УДК
60.561.9		Социология на възпитанието и образованието. Педагогическа социология	ББК
A01129/00		Машини за рязане на слама и фураж	МПК
20.23.15		Информационно-търсещи езици	ГРНТИ
29.140.50		Осветителни инсталации	МКС

Основните свойства на лексикалните единици в йерархичните класификационни езици са:

- изразът на лексикалната единица представлява условно знаково обозначение – индекс или код, а значението е фиксирано в словесен израз;
- кодовете на йерархичните класификационни езици са структурирани. В тяхната структура се отразяват йерархичните отношения между понятията. Предимство на структурираните кодове е възможността явно да бъдат изразени парадигматични

отношения, което позволява да се установи и използва връзката между понятията при търсене. Пример от УДК, който илюстрира структурата на класификационните индекси и отношенията на йерархия изразени в индекса:



539.12 Елементарни частици

539.125 Нуклеони

539.125.5 Неутрони

Според метода за определяне на лексикалните единици, езикът на йерархичните класификации се числи към типа информационни езици с контролирана лексика. Всички лексикални единици се фиксират в специално разработени речници – класификационни таблици, рубрикатори, класификатори и т.н.

Парадигматични отношения в йерархичните класификации се изразяват чрез: йерархичния принцип на изграждане на таблиците, логическата структура на кодовете и чрез справочния им апарат. Йерархичният принцип на изграждане на класификационните таблици предполага отношения между понятията – подчинение и съподчинение. Структура на всеки код показва броя на йерархичните нива на класификацията. Всяко следващо ниво удължава кода, като всеки знак добавен отдясно е субординиран и отразява основните отношения в класификационните езици – подчинение (вертикал) и съподчинение (хоризонтал), както показва примерът от УДК:

Освен чрез структурата на индекса, парадигматичните отношения в йерархичните класификации се отразяват и с помощта на справочния апарат (системата от препратки и бележки). Препратките служат като средство за разграничаване и свързване на понятия с припокриващ се смисъл и за отразяване на полийерархични връзки между разделите.

Изричното изразяване на парадигматични отношения, както в структурата на класификационните индекси, така и с помощта на справочния апарат, е едно от важните предимства на йерархичните класификации. Те правят търсенето лесно и удобно като се

следва логиката на йерархията – от общото към частното, и обратно – от частното към общото. При получаване на незадоволителни резултати при търсене информационното запитване лесно може да бъде коригирано в посока разширяване или стесняване на обхвата.

Синтагматичните отношения в йерархичните класификационни езици повишават изразителността на езика, неговата семантична сила. Те позволяват по различен начин да се комбинират отделни лексикални единици (понятия), за да се изразят по-специфични понятия. Граматическите средства в йерархичните класификационни езици обикновено се изразяват със символи и знаци като: плюс, двоеточие, коса черта, апостроф, квадратни и кръгли скоби, дефис, точка и др. Те позволяват да се изразят чрез комбиниране съвсем тесни и специфични понятия, например: Приложение на химията и химическата технология в автомобилостроенето в САЩ и Канада се отразява чрез УДК така: [54+66]:629.33(71+73), където знаците плюс, квадратни скоби и двоеточие са граматически средства на информационния език. Най-сложната и развита система от граматически правила и средства сред съществуващите йерархични класификации притежава именно УДК.

Най-типичният за йерархичните класификации начин за изразяване на синтагматичните отношения е позиционната граматика, която предполага строг ред в последователността на лексикалните единици при изграждането на търсеция образ на документа или запитването. В съответствие с правилата на позиционната граматика на първо място се поставя основният индекс, към него се добавят спомагателните индекси, и, ако те са повече от един, първо се поставят индексите на специалните типови деления (специални определители), след това – общите типови деления (общи определители), чиято последователност също е регламентирана. Този определен ред при изграждане на сложни индекси, които да отразят специфични понятия се нарича класификационна формула в класификационните езици. Тя гарантира унифицираност на индексирването.

Организационната структура на съществуващите йерархични класификации е сравнително еднаква - тя включва следните компоненти: увод, основни таблици, помощни таблици и азбучно-предметен показалец (УДК, ДКД, ББК).

а) *Уводът* или аналогичните по предназначение структурни елементи обикновено представя състава и структурата, принципите на изграждане, целите, задачите и предназначението на езика. Освен това може да включва методиката на индексирване.

б) *Основните таблици* съдържат класификационни деления по области на знанието. Това е систематичен списък на всички раздели на класификацията, изградени въз основа на принципите на подчинението и съподчинението.

в) *Помощните таблици* включват класификационни деления по спомагателни признаци като език, място, време, форма и др. Помощните таблици включват понятия, общи за всички или за много от разделите (наречени общи определители) или пък се използват в отделен раздел (специални определители). В общия случай те се добавят към основните индекси, за да уточнят тяхното съдържание или формален признак (език и форма).

Наличието или отсъствието на спомагателни таблици е важна характеристика за всяка йерархична класификация. Това определя делението на йерархичните класификационни езици на изброителни и комбинационни. Изброителната класификация представлява йерархичен ред от понятия, които са разделени на отдели (класове), но повтарящите (типовите) понятия не са обособени в специални таблици, например МПК, ГРНТИ. Когато се индексира с помощта на такава класификация, документът или запитването се отнасят до адекватно деление с предварително определен индекс. Комбинационната класификация е класификационна система, в която основната класификационна таблица е допълнена от спомагателни таблици (таблици с типови деления) и класификационните индекси се образуват чрез комбиниране (ако е необходимо) на класификационните кодове на основната таблица с кодовете на типовете деления. Примери за комбинационни йерархични класификации са УДК, ББК и др.

Азбучно-предметният показалец на йерархичната класификация представлява азбучно подреден списък на всички понятия, които са представени в основните таблици и в спомагателните таблици в систематичен ред. Азбучно-предметният показалец служи като справочник на таблиците като изпълнява две основни функции: а) позволява бързо да се открие местоположението на понятието по словесна нотация в таблиците и неговия класификационния индекс; б) дава възможност да се съберат на едно място понятия, разгледани от различни аспекти, и които се намират в различни раздели на класификацията.

Основните структурни елементи на азбучно-предметният показалец са предметните рубрики. Те обикновено са три вида: прости – отразяващи понятия, които са представени на едно място в класификационната таблица; сложни – отразяват понятия разгледани в определен аспект; гнеzdови, отразяващи понятия с два или повече аспекта на разглеждане.

Примери за предметни рубрики от азбучно-предметния показалец на УДК:

Предметна рубрика	Пример
Проста	Кино 791
Сложна	Обществени сгради - архитектура 725
Гнеzdова	Корунд - като абразив 553.65 - находища 553.82 - химия 549.517.1

Конкретни йерархични класификации използвани в библиотечно-библиографската практика са представени от М. Младенова²⁴¹.

Към класификационните езици се числят и нейерархични класификации. Те се обособяват в отделна група, защото имат коренно различна структура и принципи на изграждане.

Нейерархичните класификационни езици обхващат езика на предметните рубрики и фасетните класификации.

Езикът на предметните рубрики (предметизационен език, азбучно-предметна класификация, предметна класификация, предметизационна схема) представлява нейерархична класификация на класове (предмети), разположени по азбучен ред на наименованията. Предметизационният език е основан на използването на ограничен брой предметни класове, както и при йерархичните класификации, но класовете са изразени с думи и словосъчетания, наречени предметни рубрики.

Азбуката на езика на предметните рубрики съвпада с азбуката на естествения език, на който е създадена класификацията. Освен букви от естествени езици, той може да включва цифри и пунктуационни знаци (най-често точка, запетая, тире и скоби).

Лексика. Според метода за образуване на лексикалните единици, предметизационният език е от типа словесни информационни езици. Лексикалните единици в езика на предметните рубрики се вземат от естествения език – думи, словосъчетания, абривиатури, символи, дати и др.

Предметната рубрика е кратка словесна формулировка, която отразява основния предмет на документа (или информационното запитване), аспекта в който той е разгледан и формална характеристика на документа (предназначение, форма, жанр)²⁴². Структурно предметната рубрика може да се състои от рубрика (заглавие) и подрубрика/и. Рубриката на предметната рубрика е първият елемент в многочленната предметна рубрика, който се отделя от следващия с разделителен знак (обикновено тире, в някои случаи запетая). Подрубрика на предметната рубрика е вторият и всеки следващ елемент в многочленната предметна рубрика, който се отделя от другите с разделителен знак.

Според това предметните рубрики биват едночленни (прости) и многочленни (сложни). Простата предметна рубрика се състои само от заглавие. То може да бъде от една, две или повече думи, например Танкове, Наказателно право, Хармония (музика). Сложната предметна рубрика се състои от заглавие и подрубрика/и, които определят аспектите, в

²⁴¹Младенова, Мария. Цит. съч.

²⁴²Атанасова, В. Цит. съч., с. 14.

които е разгледан предметът, или формата и предназначението на документа, например Насекоми – физиология – справочници. Сложната предметна рубрика може да бъде сложна многочленна (рубрика и подрубрика/и) или сложна описателна. Сложната описателна предметна рубрика представлява цяла фраза, която дава названието на предмета и задно с него аспекта, в който предметът се разглежда, например Децата като художници, Жените в изкуството. Както се вижда от примерите, при описателните предметни рубрики предметът и aspectът, в който се разглежда предметът, се свързват с предлог, а не се дават като рубрика и подрубрика, отделени една от друга с тире.

Парадигматичните отношения в езика на предметните рубрики се изразяват експлицитно чрез справочния апарат (система от препратки и бележки) в речника на езика. Препратките според вида на отношенията, които обозначават могат да бъдат:

– Препратка „*Виж*” отразява връзка на еквивалентност (синонимия). Чрез нея се елиминира синонимията. Тя свързва неприет термин или формулировка на предметна рубрика и приет, когато за едно понятие съществуват два или повече сходни по значение термини.

– Препратката „*Виж и*” показва йерархична връзка между предметни рубрики на принципа „от общото към частното” (род-вид). Тя може да бъде реципрочна и да свързва „от частно към по-общо” чрез препратка „*Виж и от*”.

Някои от съществуващите предметни класификации маркират и асоциативните отношения. Те могат се изразят, както чрез свързващи, така и чрез кръстосани препратки.

Многозначността (омонимията и полисемията) в предметизационния език се отстранява чрез специален синтактичен способ. Към формулировката на предметната рубрика се добавя определител – уточняваща дума поставена в скоби. Определителят не представлява независима лексикална единица и не се използва самостоятелно. Примери за предметни рубрики, при които се налага използването на определител: Хармония (музика), Треска (медицина), Основи (химия).

Синтагматичните отношения в езика на предметните рубрики се изразяват чрез морфологични и синтактични правила. Сред морфологичните правила за формулиране на предметни рубрики най-важно е правилото за избора на първата, водещата дума в предметната рубрика. Водещата дума има особено важно значение. Тя има комплексираща функция, помага да се групират близки по съдържание рубрики и да се представят рубриците в азбучен ред. Според това правило в някои случаи се прибегва до инверсия. Инверсията е условен метод, чрез който естественият ред на думите в предметната рубрика (и/или подрубрика) се размества, като на първо място се поставя тази от тях, която носи

логичното ударение, т.нар. водеща дума, например Мостове, висящи. Чрез водещата дума могат да се създадат по-широки предметни комплекси, например:

Мостове, временни

Мостове, висящи

Мостове, военни

Мостове, дървени

Мостове, каменни

Мостове, масивни

Мостове, подвижни

Мостове, пътни

Мостове, стоманобетонни

Синтактични правила се използват, когато се съставят сложни предметни рубрики. Те регламентират комбинирането на рубрика и подрубрики, с цел да се изразят по-специфични понятия. Най-важните правила гласят:

– съществуват две основни схеми на представяне на сложни предметни рубрики: Предмет – Аспект (например Сърце – изследване) и Предмет – Показател на връзката предмет (например Пептиди – влияние върху организма);

– в сложна предметна рубрика не може да се свързват лексикални единици, в отношения на подчинение и противопоставяне;

– в сложната предметна рубрика елементите трябва да бъдат така разположени, че при превода на рубриката на естествен език да не се наруши смисъла на изходното значение;

– обикновено в многочленна предметна рубрика отделните елементи следват реда: рубрика, тематична подрубрика, географска подрубрика, формална подрубрика.

Тези правила се основават на позиционната граматика. Използването на тематични, хронологични, географски и формални подрубрики, които изпълняват в структурата на търсещия образ същите функции, както и общите определители за място, време и форма в УДК, може да се разглежда като прилагане на граматичния метод „показатели за роля”.

За да се извършва индексирание на документи и запитвания чрез на езика на предметните рубрики е необходим нормативен речник - списък на предметните рубрики като средство за контрол на лексиката. Списъкът на предметните рубрики е азбучна ред на лексикалните единици, използвани като предметни рубрики или подрубрики с установените между тях парадигматични отношения.

Много от съществуващите предметни класификации включват някои характеристики свързани със структурата на тезаурусите. Например отразяването на родово-видовите отношения и асоциативните връзки между термините е особеност на тезауруса²⁴³.

3.3.3. Информационни езици за многоаспектно представяне на съдържанието на документите

За многоаспектно представяне на съдържанието на документите са предназначени посткоординатните езици. За разлика от класификационните езици (от предкоординатен тип), при които предметът (темата) на документа се изразява общо и компактно чрез комбиниране на термини в момента на индексиранието на документа²⁴⁴, посткоординатните езици използват за отразяване на документа множество думи (практически без ограничения на техния брой) от текста, заглавието и/или резюмето на документа или свързани с тях. Те са основани на принципа на посткоординацията – свързването на термините, които представят съдържанието на документа става в момента на търсенето²⁴⁵.

Посткоординатните езици са от словесен тип, използват за изразяване на лексикалните си единици естествен език. Последователността на лексикалните единици в търсещия образ на документа или на запитването не е от особено значение, за разлика от класификационните езици, при които са в сила строгите правила на позиционната граматика и редът на свързване на отделни лексикални единици е точно определен.

Създаването и използването на посткоординатни езици цели да се избегнат някои съществени ограничения на класификационните езици като сложност при индексиранието (процесът не може да бъде автоматизиран напълно), сложени синтактични правила, които са неразбираеми за неопитни потребители, нужда от квалифицирани индексатори.

Посткоординатните езици обхващат езика на ключовите думи и дескрипторен език.

Езикът на ключовите думи предполага използване на думи и устойчиви словосъчетания, взети от текста на документа, заглавието и/или резюмето. Лексикалните единици, наречени ключови думи включват само смислово натоварени думи (без съюзи,

²⁴³ ИФЛА. Руководство по обеспечению тематического доступа..., с. 23.

²⁴⁴ Markey, Karen. Forty Years of Classification Online: Final Chapter or Future Unlimited?. In: *Cataloguing and Classification Quarterly*, 2006. Vol. 42, № 3, p. 1–63.

²⁴⁵ Svenonius, Elaine. Precoordination or not?. In: *Subject indexing: principles and practices in the 90's: Proceedings of the IFLA Satellite meeting held in Lisbon, Portugal, 17–18 Aug. 1993*. München: Saur 1995, p. 231–255.

предлози и др.п.). Обикновено ключовите думи се привеждат в стандартна лексикографска форма (число, род и др.).

Лексиката на езика на ключовите думи е неконтролирана. Това предполага простата при търсене и скорост на индексирание, възможно и напълно автоматично, но се свързва и със съществени недостатъци: ниска степен на точност (възможност да възникне грешна координация) и липса на контекстуалност (неясни отношения между термините).

Дескрипторният език компресира тези ограничения на езика на ключовите думи. Неговите лексикални единици – дескрипторите, представляват смислова доминанта на клас условно еквивалентни понятия (синоними) в рамките на конкретен дескрипторен език²⁴⁶. Дескрипторите, за разлика от ключовите думи, подлежат на лексикографска и семантична обработка за отстраняване на синонимията и многозначността, и за установяване на парадигматичните отношения. Дескрипторите се представят в контролиран речник, наречен тезаурус.

Тезаурусът представлява речник на обикновено тематично ограничено множество от понятия с обозначение на характерните за тях семантични връзки²⁴⁷. Създаването на тезаурус е свързано с функционирането на конкретна информационна система. Той служи за терминологичен контрол при въвеждането и търсенето на информация и като ръководство за превод на съдържанието на документите от естествен на информационен език, за допълнително индексирание на документите и информационните запитвания.

Най-актуалните стандарти са съобразени с концепцията на семантичния уеб за представянето на тезаурусите като свързани данни в уеб*. Международните стандарти са приети като национални стандарти и в България (БДС ISO 25964-1:2012 Информация и документация. Тезаурус и способност за взаимодействие с други речници и БДС ISO 25964-2:2016 Информация и документация. Тезаурус и взаимодействие с други речници. Част 2: Взаимодействие с други речници и на английски език). Стандартите дават указания за

²⁴⁶ Габровска, Свободозаря, Пейчева, Веселина. Цит. съч., с. 29.

²⁴⁷ Пак там, с. 30.

*ISO 25964-1:2011 Information and documentation. Thesauri and interoperability with other vocabularies. Part 1: Thesauri for information retrieval, Information and documentation; ISO 25964-2:2013. Information and documentation. Thesauri and interoperability with other vocabularies. Part 2: Interoperability with other vocabularies; ANSI/NISO Z39.19-2005 (R2010) Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies; BS 8723-2:2005 Structured vocabularies for information retrieval. Guide. Thesauri; BS 8723-1:2005. Structured vocabularies for information retrieval. Guide. Definitions, symbols and abbreviations, BS 8723-3:2007 Structured vocabularies for information retrieval. Guide. Vocabularies other than thesauri.

приложението на тезаурусите, тяхната лексика, семантични отношения и структура и кратка методика за тяхното съставяне и редактиране.

Съставянето на тезаурус в конкретна тематична област се разглежда от различни изследователи още началния им етап на развитие. На български език съществуват изследвания, посветени на съставянето на тезауруси в обществените науки²⁴⁸.

Изследователите обръщат внимание на тенденцията на сближаване на тезауруса и списъка на предметните рубрики като инструменти за индексирание²⁴⁹, макар че списъкът на предметни рубрики води своето начало от библиотечните традиционни каталози и е предназначен за индексирание на библиотечен фонд с универсална тематика, а тезаурусът от появата му се използва в бази от данни в автоматизиран режим и е предназначен за посткоординатно индексирание най-вече за научни документи (статии, доклади, дисертации) в специфични предметни области. Въпреки това в съвременните лингвистични инструменти – списъци на предметни рубрики и тезауруси, разликите все повече се заличават. Много от предметизационните класификации използват прийоми на тезаурусите – отразяване на семантични връзки между термини с широко значение (*broader term*) и тясно значение (*narrower term*), и свързани термини (*related term*). От друга страна някои по-нови тезауруси се разработват като инструменти за индексирание на документален масив с универсална тематика, например италианският общ тезаурус „Nuovo Soggettario”, който се използва за тематичен достъп до националната библиография на Италия*.

На база представянето на видовете информационните езици, с оглед ролята им в лингвистичното осигуряване на електронна библиотека бяха изведени някои основни предимства и ограничения по конкретно на езиците, които отразяват съдържанието на документите (Приложение 3). Те ще бъдат взети под внимание при избора на информационни езици в състава на лингвистичното осигуряване на българска електронна библиотека с универсално съдържание.

3.4. Анализ на резултатите от изследване на използвани лингвистични средства в българските електронни библиотеки.

Функциите на електронната библиотека, свързани с достъпа до документите - търсене, намиране, навигиране, представяне и получаване на документи се реализират чрез

²⁴⁸ Габровска, Свободозаря, Пейчева, Веселина. Цит. съч., с. 29–60.

²⁴⁹ Атанасова, В. Цит. съч., с. 3.

* <http://thes.bncf.firenze.sbn.it/>

различни информационни езици, които в съвкупност съставят лингвистичното осигуряване на електронната библиотека.

Информационните езици изпълняват общети функции: търсене/навигиране, преобразуване, структуриране, систематизиране и идентифициране. По-конкретно всеки отделен език, съобразно своята специфика може изпълнява различни функции. Езиците в лингвистичното осигуряване на електронна библиотека с техните типологизарици признаци и специфични функции са представени в табл. 10.

Таблица 10. Информационни езици: видове и функции.

Информационни езици	Тип и вид на информационния език	Функция
Език на библиографското описание	словесен информационен език	Идентифициране Преобразуване Структуриране Търсене/Навигиране
Език на йерархичните класификации	предкоординатен, кодов информационен език, контролиран	Структуриране Систематизиране Преобразуване Търсене/Навигиране
Език на предметните класификации	предкоординатен, словесен информационен език, контролиран	Преобразуване Структуриране Търсене/Навигиране
Дескрипторен език	посткоординатен, словесен информационен език, контролиран	Преобразуване Структуриране Търсене
Език на ключовите думи	посткоординатен, словесен информационен език, неконтролиран	Търсене

Използването на различни информационни езици предоставя различни възможности за търсене в електронната библиотека. Търсенето се разглежда като активен процес за откриване на документ/и по различни признаци. Обикновено в електронна библиотека се осигуряват следните видове търсене:

– търсене по дума или словосъчетание в пълния текст и/или резюмето на документ/и, определяно като лексикално търсене (пълнотекстово търсене). Този метод предполага използване на математически и статистически методи за търсене и ранжиране на резултатите и не е обект на нашето изследване;

– търсене по формални признаци на документа, например име на автор, заглавие, дата на публикуване и т.н., наричано библиографско търсене. Библиографското търсене се основава на общ модел за представяне на признаците (атрибутите) на документите, т.е. по елементите от данни на библиографското описание;

– навигирането е логическото свързване и представяне на библиографските данни по общи признаци, например всички документи по една и съща тема, по вид на документите и т.н.;

– търсене на документи по тяхното съдържание – тематично търсене. То се основава на езиците за отразяване на съдържанието на документите:

⇒ Класификационните езици съотнасят всеки документ към предметни класове, предварително фиксирани в някаква класификационна схема (йерархична или предметна класификация). Най-същественото предимство на този подход е, че осигурява структуриране и систематизиране по тема в електронната библиотека. Затова използването на класификация е задължително, особено в универсални и широко тематични електронни библиотеки.

⇒ Словесните посткоординатни езици отразяват съдържанието на документа, чрез думи от самия документ (текста, заглавието, резюмето), както е при езика на ключовите думи, и чрез думи, смислово и логически свързани с ключови думи от текста (дескрипторен език).

I. Характеристики на изследването

1. Цели на изследването

Изследването се провежда със следните цели:

1.1. Да се проучи опитът на българските електронни библиотеки в използването на информационни езици за представяне на характеристиките и съдържанието на документите с цел достъп, търсене и навигиране.

1.2. Прилагане на резултатите от анализа на изследването в модел на лингвистично осигуряване на българска електронна библиотека с универсално съдържание.

2. Период на провеждане:

Изследването се проведе в периода 1–30 май 2018 г. и обхваща общо 24 електронни библиотеки в интернет (от предварително идентифицираните 38). Отпадането на някои от тях се дължи на неотговаряне на някои от критериите (т. 5) и/или затова, че не са били

достъпни в момента на изследването (Дигиталната библиотека на Българската Патриаршия, Дигитална библиотека на Регионална библиотека „Ст. Чилингиров” – Шумен).

3. Изследователска методология

За реализиране на целите на изследването са използвани следните методи:

3.1. Директно проучване, което предполага провеждане на многократни сесии за търсене с използване на наличните информационни езици и/или методи и средства за търсене във всяка от подбраните за изследване електронни библиотеки.

3.2. Кабинетно проучване (Desk research) – методът е приложен главно при анализа на приложението на информационните езици.

4. Реализация на изследването

Изследването е проведено в българското интернет пространство върху български електронни библиотеки, предоставени на безплатен достъп и без регистрация.

5. Критерии за подбор

За подбора на български електронни библиотеки, които да бъдат включени в изследването бяха приложени следните критерии, изведени въз основа на приетата в настоящия труд работна дефиниция:

5.1. Онлайн ресурси, представени в българското интернет пространство на български език, които са назовани „електронна библиотека”, „дигитална библиотека”, „виртуална библиотека”, „онлайн библиотека”, „цифрова библиотека”;

5.2. Съдържат пълнотекстови документи;

5.3. Предоставят услуги (функции) – достъп, търсене и/или навигиране.

6. Задачи на изследването

Изпълнението на изследването предполага реализиране на няколко задачи:

6.1. Търсене в българското интернет пространство на електронни ресурси, назовани „електронна библиотека”, „дигитална библиотека”, „виртуална библиотека”, „онлайн библиотека”, „цифрова библиотека”. Търсенето бе проведено в заглавията на ресурсите чрез търсачката Google като резултатите бяха ограничени по държава и език („България” и „български”). Проведени бяха няколко сесии на търсене по избраните ключови думи.

6.2. Преглеждане на получените резултати. Получиха бяха общо 38 резултата, релевантни за търсенето по ключови думи „електронна библиотека”, „дигитална библиотека”, „виртуална библиотека”, „онлайн библиотека”.

6.3. Провеждане на вторичен подбор по критерии 5.2. и 5.3. В резултат от прилагането на критериите бяха селектирани 24 електронни библиотеки, които предоставят

достъп до пълнотекстови документи (различни по тип/вид и тематика) и предлагат услуги на потребителите – достъп, търсене и/или навигиране.

6.4. Групиране на изследваните онлайн ресурси по признака принадлежност с цел извеждане на зависимости, свързани с опита на традиционните библиотеки в разработването и използването на информационни езици. От избраните 24 електронни библиотеки за изследване на езиците в лингвистичното осигуряване бяха формирани следните групи:

- електронни библиотеки, създадени от физически лица*;
- електронни библиотеки на обществени организации и учреждения;
- електронни библиотеки на научни институти;
- електронни библиотеки на университети;
- електронни библиотеки, създадени от обществени библиотеки.

6.5. Провеждане на директно проучване на използваните информационни езици в избраните електронни библиотеки, което обхваща:

- установяване на методи и средства на търсене (основно търсене, разширено търсене, използване на логически оператори, контекстни оператори и др.);
- установяване по какви признаци (характеристики) на документите се провежда търсене (полета, елементи от данни);
- установяване на какви информационни езици се базира търсенето.

6.6. Представяне и обобщаване на събраните данни.

В резултат на изпълнението на задачите на изследването бяха събрани и обобщени данни за 24 български електронни библиотеки, отговарящи на критериите: да са назовани „електронна библиотека”, „дигитална библиотека”, „виртуална библиотека”, „онлайн библиотека”, „цифрова библиотека”; да съдържат пълни текстове; да предлагат функции достъп, търсене и/или навигиране. Електронните библиотеки са групирани според принадлежност: 4 електронни библиотеки на физически лица; 4 електронни библиотеки на обществени организации и учреждения; 4 електронни библиотеки на научни институти на БАН; 5 електронни библиотеки на университети; 7 електронни библиотеки, създадени от обществени библиотеки. Данните се отнасят за използването на информационни езици за достъп, търсене и/или навигиране, както и за методи и средства за търсене в електронните библиотеки и са представени в табл. 11.

* включително сдружения на лица, например Българска виртуална библиотека „Словото” се управлява от Сдружение с идеална цел „Словото”.

Таблица 11. Български електронни библиотеки.

1. Електронни библиотеки на физически лица (4)				
Название	Уеб адрес	Съдържание	Търсене/Навигиране	Информационни езици
1. Електронна библиотека „Моята библиотека”	https://chitanka.info/	Художествена литература; Научнопопулярна литература; Научна литература.	Търсене в едно поле по данни от библиографското описание. Навигиране по азбучен ред на авторите и по собствена класификационна схема по тема	Език на библиографското описание; Предметна класификация.
2. Българска виртуална библиотека „Словото”	https://www.slovo.bg/	Българска художествена литература.	Търсене в едно поле по данни от библиографското описание (автор и заглавие, включително по заглавие на структурни части на документа, напр. глави). Навигиране по имената на колекциите и по азбучен ред на авторите.	Език на библиографското описание.
3. Виртуална библиотека „Спиралата на еволюцията”	http://www.spiralata.net/biblioteka/	Научнопопулярна литература.	Търсене в едно поле по данни от библиографското описание (автор и заглавие). Навигиране по азбучен ред на имената на авторите и по азбучен ред на заглавията.	Език на библиографското описание.
4. Електронна библиотека по архивистика и документалистика	http://www.electronic-library.org/content/	Научна литература; Учебна литература.	Търсене в едно поле по данни от библиографското описание и думи от пълния текст. Навигиране по имена на колекциите, а вътре в тях по вид на документа – монографии, статии;	Език на библиографското описание.

			Навигиране по азбучен ред на авторите в две поредици български автори и чужди автори.	
2. Електронни библиотеки на обществени организации и учреждения (4)				
Название	Уеб адрес	Съдържание	Търсене/Навигиране	Информационни езици
5. Електронна библиотека на Център за развитие на човешките ресурси и регионални инициативи към Министерство на труда и социалните грижи	http://www.e-libraryeu.com/	Специализирана литература.	Търсене в едно поле по автор и заглавие. Навигиране по имената на колекциите.	Език на библиографското описание.
6. Електронна библиотека на Институт Сервантес	http://sofia.cervantes.es/bg/biblioteca_ispanski/electronna_biblioteka_ispanski.htm http://catalogo-bibliotecas.cervantes.es/general/abnetcl.exe/O7141/ID2567138d/NT1?ACC=111&LANG=en-US	Художествена литература от испански и латиноамерикански автори – е-книги и аудио книги; Справочна литература; Научна литература от областите лингвистика, изкуство, философия и др.; Материали за изучаване на испански език.	Търсене в едно поле по библиографски данни и разширено търсене с комбиниране на признаци: автор, заглавие, предметни рубрики, издателство, имена колекция, година издаване, език, страна, формат.	Език на библиографското описание; Език на предметните рубрики. <i>Бележка:</i> <i>електронната библиотека е интегрирана с електронен каталог.</i>
7. Дигитална Библиотека на Национален Статистически Институт	http://statlib.nsi.bg:8181/bg/index.php	Специализирана литература.	Търсене с възможност за съчетаване по признаци: автор/колективен автор, заглавие, заглавие на периодично издание, предметни рубрики, серия, година на издаване, местоиздаване,	Език на библиографското описание. Език на предметните рубрики.

			издателство, ISBN/ISSN, език, по дума/и от библиографското описание.	<i>Бележка: електронната библиотека е интегрирана с електронен каталог.</i>
8. Виртуална библиотека на Алеф Център за еврейско-българско сътрудничество	http://alef-bg.org/virtualna-biblioteka/	Документи (вкл. видео и интернет публикации), свързани с еврейско-българското сътрудничество.	Търсене в едно поле по автор и заглавие. Навигиране по имената на колекциите.	Език на библиографското описание.
3. Електронни библиотеки на научни институти (4)				
Название	Уеб адрес	Съдържание	Търсене/Навигиране	Информационни езици
9. Електронна библиотека за тракийско изкуство Институт за изследване на изкуствата, БАН	http://thracian-art.artstudies.bg/?p=library	Научни публикации; Артефакти: живопис, монети, скулптура, предметни форми.	Търсене в едно поле по автор, заглавие, ключови думи, които отразяват съдържанието на текстовите документи. Разширено търсене във всяка от колекциите: Живопис, Монети, Скулптура, Предметни форми по метаданни на артефактите (с комбиниране на признаците за търсене).	Език на библиографското описание. Език на ключовите думи.
10. Виртуална библиотека „Иван Шишманов“: Българската литература в превод Институт за литература, БАН	http://bglit.org/page/bg/nachalo.php	Преводи на българска художествени произведения на немски и английски езици	Търсене с възможност за съчетаване на признаците: автор, заглавие, преводач, редактор/съставител, година на издаване, език, ключови думи, които отразяват съдържанието на документите.	Език на библиографското описание; Език на ключовите думи.

<p>11. Електронна библиотека „Българско езикознание” на Института за български език „Проф. Любомир Андрейчин”, БАН</p>	<p>http://ibl.bas.bg/informatsiya/uslugi/elektronna-biblioteka/</p>	<p>Публикации на Института за български език „Проф. Любомир Андрейчин”</p>	<p>Търсене в едно поле по думи от библиографското описание. Представяне на документите в списък.</p>	<p>Език на библиографското описание.</p>
<p>12. Виртуална библиотека на Институт по биофизика, Институт молекулярна биология, Институт по физиология, БАН</p>	<p>http://www.bio21.bas.bg/library/</p>	<p>Специализирана научна литература.</p>	<p>Търсене с комбиниране по признаците: тип на документа, заглавие, автор, година на публикуване, думи от заглавието.</p>	<p>Език на библиографското описание.</p>
<p>4. Електронни библиотеки на университети (5)</p>				
<p>Название</p>	<p>Уеб адрес</p>	<p>Съдържание</p>	<p>Търсене/Навигиране</p>	<p>Информационни езици</p>
<p>13. Дигитална библиотека на Нов български университет</p>	<p>https://nbu.bg/bg/librariy/elektronni-resursi/digitalna-biblioteka</p>	<p>Научна литература; Учебна литература.</p>	<p>Търсене в едно поле по думи от библиографското описание. Навигиране по автор, заглавие, година на издаване, УДК индекс.</p>	<p>Език на библиографското описание; Език на йерархичните класификации. <i>Бележка:</i> <i>електронната библиотека е интегрирана с НАБИС репозиториум.</i></p>

<p>14. Дигитална библиотека Университет за национално и световно стопанство</p>	<p>http://www.unwe.bg/library/bg/</p>	<p>Специализирана научна и учебна литература</p>	<p>Търсене с възможност за съчетаване на признаци: автор, заглавие, година на издаване, издателство, заглавие на периодично издание, ключови думи, които отразяват съдържанието на документа, УДК индекс.</p>	<p>Език на библиографското описание; Език на ключовите думи; Език на йерархичните класификации.</p>
<p>15. Дигитална библиотека СУ "Св. Климент Охридски" SUDigital</p>	<p>http://lib.sudigital.org/?ln=bg</p>	<p>Разнообразни по тип медия, вид и тематика документи: архивни, документи, ръкописни книги, живопис, фотографии, карти, периодични издания, учебни ресурси, статии, монографии, публикации на преподавателите от Исторически факултет и др.</p>	<p>Търсене в едно поле по признаци: заглавие, автор, резюме, ключови думи, пълен текст и разширено търсене с възможност за съчетаване на признаци [<i>търсенето не работи</i>]. Преглеждане на документите по имената на колекциите.</p>	<p>Език на библиографското описание.</p>
<p>16. Зографска електронна научно-изследователска библиотека (СУ - Катедра „Кирилометодиевистика“, Център за славяно-византийски проучвания „Проф. Иван Дуйчев“ Университетска библиотека)</p>	<p>https://zograflib.slav.uni-sofia.bg/</p>	<p>Славянската ръкописна сбирка на Зографската Света обител и други колекции от славянски ръкописи</p>	<p>Търсене в едно поле по данни от библиографското описание. Навигиране по индекс на: ключови думи, които отразяват съдържанието на документите; на автори; на заглавията</p>	<p>Език на библиографското описание; Език на ключовите думи.</p>
<p>17. Дигитална библиотека Аграрен университет - Пловдив</p>	<p>http://lib.auplovdiv.bg/dglb/</p>	<p>Специализирана научна литература, Периодични издания до 1944 г.</p>	<p>Търсене няма. Навигиране по азбучен ред на имената на авторите, по азбучен ред на заглавия, по</p>	<p>Език на библиографското описание;</p>

			азбучен ред на предметни рубрика, по азбучен ред на ключови думи.	Език на предметните рубрика; Език на ключовите думи.
5. Електронни библиотеки на обществени библиотеки (7)				
Название	Уеб адрес	Съдържание	Търсене/Навигиране	Информационни езици
18. Варненска дигитална библиотека Регионална библиотека "Пенчо Славейков"	http://digitallibrary.library.bg/	Периодични издания, пощенски картички и фотографии от края на XIX и началото на XX век, свързани с гр. Варна	Търсене в едно поле по признаци: заглавие, думи от анотация, имена на автори/лица, предметна рубрика, географска рубрика.	Език на библиографското описание; Език на предметните рубрики.
19. Дигитална библиотека на Народна библиотека "П. Р. Славейков"	http://www.libraryvt.com/дигитална-библиотека.html	Научнопопулярни издания; Художествена литература; Научни публикации с библиотечна тематика; Библиографски издания, Издания на библиотеката; Местен периодичен печат.	Търсене няма. Навигиране по колекции.	Език на библиографското описание.
20. Дигитална библиотека Стара Загора	https://www.rodina-bg.org/digitalna-biblioteka	Местен периодичен печат; книги издадени до 1944 г.	Търсене няма. Представяне на документите в списък.	Език на библиографското описание.
21. Дигитална библиотека „Габровски периодичен печат“ Регионална библиотека „Априлов - Палаузов“ - Габрово	http://e-lib.libgabrovo.com/	Местен периодичен печат.	Търсене в едно поле по думи от библиографското описание и анотацията. Навигиране по азбучен ред на заглавията на периодичните издания.	Език на библиографското описание.

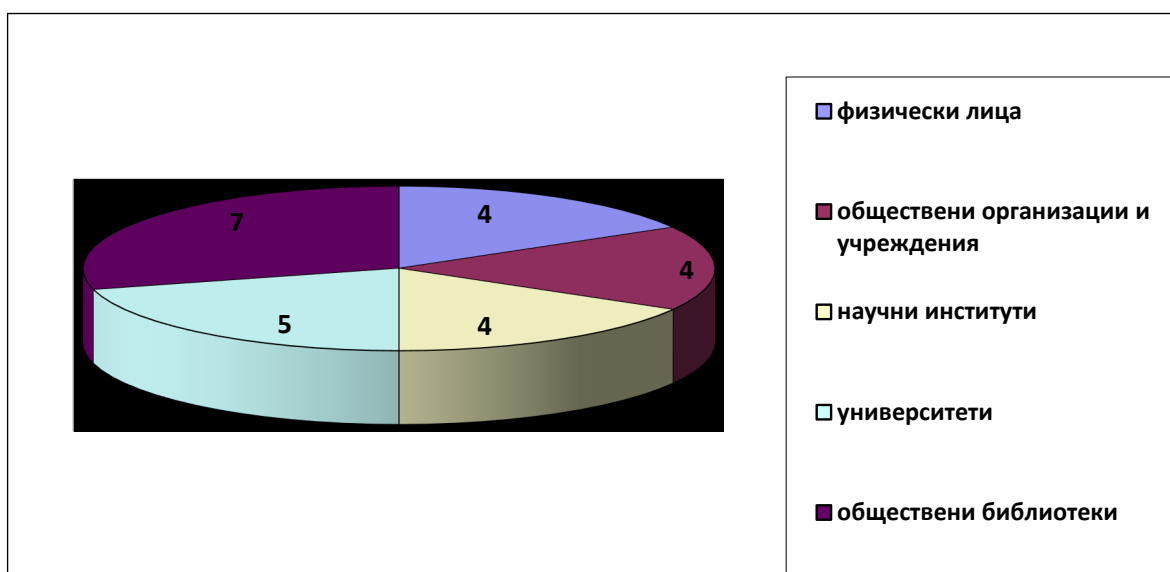
<p>22. Дигитална библиотека "Мост през вековете" Регионална библиотека „Петър Стъпов” - Търговище</p>	<p>http://libtg.info/page/bg/digitalna-biblioteka/digitalna-biblioteka-most-prez-vekovete.php</p>	<p>Местен периодичен печат.</p>	<p>Търсене в едно поле по думи от библиографското описание, анотация и ключови думи към статиите.</p>	<p>Език на библиографското описание; Език на ключовите думи.</p>
<p>23. „Дигитална библиотека Хасково” Регионална библиотека „Христо Смирненски”</p>	<p>http://rb-haskovo.no-ip.org/absw/buletin-Digital/Default.HTM</p>	<p>Пощенски картички, фотографии, архивни документи, свързани с града.</p>	<p>Търсене няма. Навигиране чрез индекси на заглавия, предметни рубрики, имена на автори и лица.</p>	<p>Език на библиографското описание; Език на предметните рубрики.</p>
<p>24. Дигитална библиотека на Национална библиотека „Св. св. Кирил и Методий”</p>	<p>http://www.nationallibrary.bg/wp/?page_id=294&lang=bg</p>	<p>Ръкописи, старопечатни книги, непубликувани документи от Българския исторически архив и от Ориенталския отдел, портрети и снимки, графични и картографски издания, български вестници и списания от периода 1844-1944 г.</p>	<p>Всяка колекция (28 на брой) предполага признаци за търсене, според вида на документите (автор, заглавие, издателско, година на издаване, тема, ключови думи и др.).</p>	<p>Език на библиографското описание. Език на ключовите думи; Език на предметните рубрики.</p>

II. Анализ на резултатите от проучването

При провеждането на изследването се установи, че в интернет пространството не съществува списък (регистър) на български електронни библиотеки и съответно не са налични критерии, чрез които да е възможно класифицирането на електронните библиотеки. Това обстоятелство наложи изработване на собствена критериална система (5.1-5.3), която обхващаща възможни индикатори, позволяващи въз основа на събраната на информация и търсенето на зависимости анализиранието и направата на изводи за състоянието на лингвистичното осигуряване на българските електронни библиотеки.

Според институционалната принадлежност се установи, че най-голям брой електронни библиотеки са създадени от традиционни (обществени) библиотеки – 7 на брой и от университети – 5. Трябва специално да бъде подчертано, че в изследването не са включени научните институционални репозиториуми, тъй като не отговарят на критерии, посочени в т. 5.1* (фиг. 20)

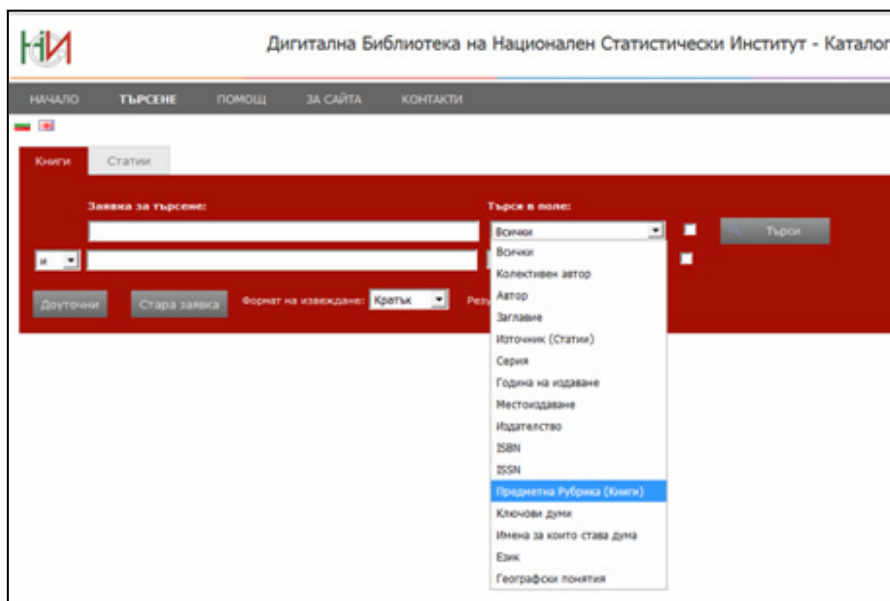
Фигура 20. Български електронни библиотеки според принадлежност



* В OpenDOAR са регистрирани общо 8 институционални репозиториума. Единият от тях е нар. „дигитална библиотека” - BulDML (Bulgarian Digital Mathematics Library at IMI-BAS), но не е включен в изследването, за да се спази общото правило да не бъдат изследвани институционални репозиториуми. Макар, че репозиториумите отговарят на широкото определение за електронна библиотека, което приемаме, те не са включени в изследването по две причини: 1. не отговарят на формалния критерий да бъдат наречени „библиотека” и 2. Както бе показано гл. 1, т. 1.5. българските репозиториуми използват 2 програмни инструмента – DSpace (6) и EPrints (2). Те предполагат определен метод на навигиране и набор от средства за търсене, което би насочило резултатите от изследването в една посока.

В някои случаи електронните библиотеки са интегрирани с електронен библиотечен каталог или репозиториум (Дигитална библиотека на Нов български университет, Дигитална Библиотека на Национален Статистически Институт, Електронната библиотека на Институт Сервантес). Те предлагат по-развити възможности за търсене по различни признаци с възможност за използване на логически оператори и използват повече информационни езици, за отразяване на съдържанието на документите (ил. 1).

Илюстрация 1. Дигитална библиотека на Национален статистически институт.



Изследването показва, че електронните библиотеки използват различни подходи за достъп до документите по съдържание. Достъпът може да бъде структурно-логически (навигиране) и чрез търсене. При структурно-логическия достъп навигирането се реализира чрез рубрики (категории) на отрасли, теми, понятия. Това може да се осъществи чрез налични класификации (например по УДК – Дигитална библиотека на Нов български университет) или чрез специално разработени за конкретната електронна библиотека списъци на тематични рубрики или формални категории (например Електронна библиотека „Моята библиотека”, виж ил. 2).

Илюстрация 2. Тематични рубрики на Електронна библиотека „Моята библиотека”.

По категория		
<ul style="list-style-type: none"> • @Некатегоризирани 2 • Наука, технологии, образование 198 <ul style="list-style-type: none"> ○ Научнопопулярна литература 57 <ul style="list-style-type: none"> ○ Образование 3 <ul style="list-style-type: none"> ▪ Дидактика ▪ Педагогика 1 ▪ Учебна литература 1 ○ Природоматематически науки 8 <ul style="list-style-type: none"> ▪ Астрономия ▪ Биология 1 ▪ География ▪ Екология ▪ Икономика и финанси 5 ▪ Информатика ▪ Математика ▪ Медицина 2 ▪ Физика ▪ Химия ○ Справочна литература 10 <ul style="list-style-type: none"> ▪ Енциклопедии и речници 3 ▪ Ръководства и самоучители 2 ▪ Справочници 4 ○ Технологии <ul style="list-style-type: none"> ▪ Високи технологии ▪ Информационни технологии ▪ Съвременни и модерни технологии ▪ Традиционни технологии ○ Хуманитарни и социални науки 120 <ul style="list-style-type: none"> ▪ Антропология и етнология 10 ▪ Етнография и фолклористика 1 ▪ История 61 ▪ Културология 6 ▪ Литературознание 12 ▪ Политология ▪ Психология 8 ▪ Семiotика ▪ Социология 1 ▪ Философия 21 • Общество и култура 151 <ul style="list-style-type: none"> ○ Домашен бит и семейство 24 <ul style="list-style-type: none"> ▪ Градинарство ▪ Грижи за детето 2 ▪ Домакинство и бит ▪ Домашни животни ▪ Здраве и алтернативна медицина 8 ▪ Кулинария 4 ▪ Приложна психология 8 ▪ Развлечения и свободно време ▪ Семейство 2 ▪ Хоби и занаяти ○ Изкуство 8 <ul style="list-style-type: none"> ▪ Архитектура 	<ul style="list-style-type: none"> • Изобразително изкуство • Кино • Музика 2 • Приложни изкуства • Театър 1 ○ Общество 36 <ul style="list-style-type: none"> ▪ Бизнес ▪ Култура 1 ▪ Политика 31 ▪ Спорт 2 ▪ Юридическа литература 2 ○ Религия и парарелигиозни учения и практики 83 <ul style="list-style-type: none"> ▪ Езотерика и астрология 6 ▪ Религия 5 ▪ Самоусъвършенстване и алтернативно познание 72 • Разкази в картинки 228 • Художествена литература 892 <ul style="list-style-type: none"> ○ Басни 3 ○ Българска литература 560 <ul style="list-style-type: none"> ▪ Българска военна проза 6 ▪ Българска документална проза 43 ▪ Българска драматургия 3 ▪ Българска историческа проза 61 ▪ Българска класика: поезия 68 ▪ Българска класика: романи и повести 23 ▪ Българска криминална проза 9 ▪ Българска публицистика 6 ▪ Българска фантастика 90 ▪ Българско фентъзи 10 ▪ Възрожденска литература 1 ▪ Старобългарска литература 1 ▪ Съвременна българска поезия 67 ▪ Съвременна българска проза: разкази и новели 63 ▪ Съвременни български романи и повести 80 ○ Детска литература 492 <ul style="list-style-type: none"> ▪ Детски пиеси и сценки ▪ Детски стихове и гатанки 30 ▪ Приказки 116 ▪ Разкази за деца 15 ▪ Романи и повести за деца 281 ○ Документална проза 130 <ul style="list-style-type: none"> ▪ Биографии и автобиографии 45 ▪ Дневници 3 ▪ Документални очерци 7 ▪ Мемоари 22 ▪ Писма 1 ▪ пътешествия и експедиции 47 ○ Драматургия 19 <ul style="list-style-type: none"> ▪ Исторически драми 2 ▪ Комедии 3 ▪ Семейни драми ▪ Социални драми 1 	<ul style="list-style-type: none"> • Сценарии • Трагедии 7 • Трагикомедии 3 ○ Исторически романи и повести 263 <ul style="list-style-type: none"> ▪ Книги игри 56 ○ Криминална проза 608 ○ Любовни романи и повести 2212 ○ Поезия 49 <ul style="list-style-type: none"> ▪ Класическа поезия 22 ▪ Съвременна поезия 18 ○ Приключенска литература 286 ○ Проза 745 <ul style="list-style-type: none"> ▪ Военна проза 20 ▪ Есета 8 ▪ Класическа проза: разкази и новели 36 ▪ Класически романи и повести 109 ▪ Съвременна проза: разкази и новели 77 ▪ Съвременни романи и повести 479 ○ Публицистика 3 ▪ Интервюта ▪ Репортажи ▪ Статии ○ Сборници с разнообразно съдържание 31 ○ Старинна литература 72 <ul style="list-style-type: none"> ▪ Антична литература 16 ▪ Древновизантична литература 10 ▪ Европейска средновековна и ренесансова литература 25 ▪ Митове, легенди, епос 21 ○ Тримър 858 ○ Фантастика 1603 <ul style="list-style-type: none"> ▪ Научна фантастика 793 ▪ Фентъзи 562 ▪ Хорър 120 ○ Фолклор 3 ○ Хумор и сатира 96

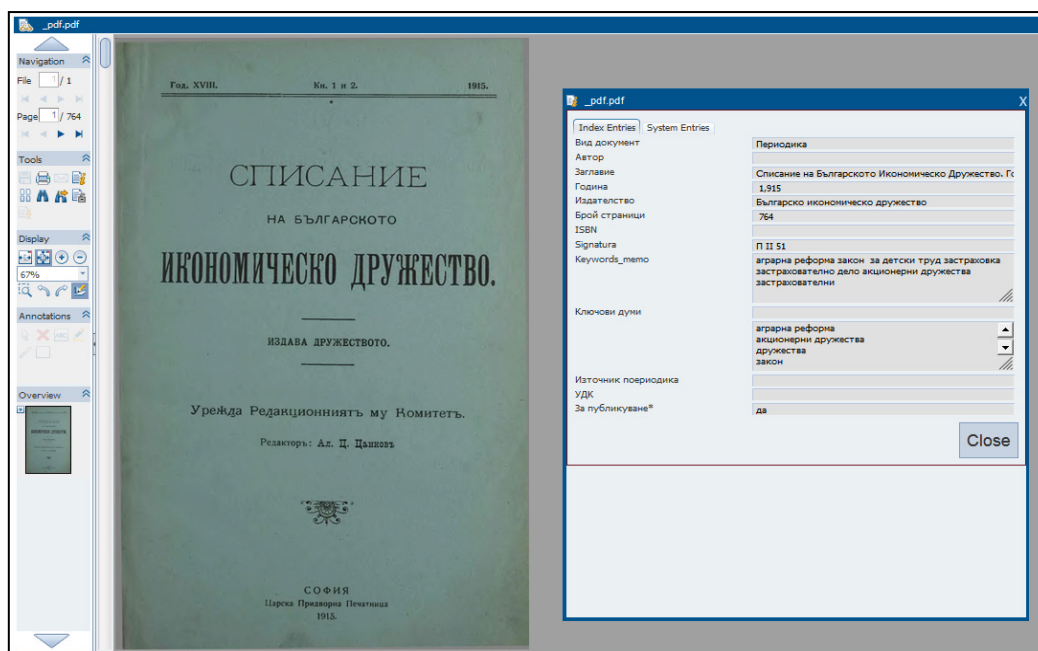
Изследваните български електронни библиотеки предлагат търсене в едно поле (основно търсене), което предполага търсене по дума/и или словосъчетание/я от библиографското описание на документа (при всички изследвани библиотеки, които предлагат функцията търсене) или дума от анотацията на документите (във Виртуална библиотека „Спиралата на еволюцията”).

Някои електронни библиотеки предвиждат и разширено търсене – търсене по няколко признаци на документа (автор, заглавие или др. елементи от библиографското описание) чрез комбиниране с булеви оператори.

Най-често изследваните електронни библиотеки предлагат търсене по ключови думи. Наблюдението показва, че под търсене по ключови думи, разработчиците на електронните библиотеки разбират съвършено различни методи: а) специално подбрани думи от текста (заглавието и/или резюмето) на документа, които отразяват съдържанието на документите (език на ключовите думи); б) думи от заглавието на документа; в) думи от библиографското описание; г) всички думи от анотацията/резюмето на документа или

от текста на документа. В общия случай това не е уточнено за потребителя. Езикът на ключовите думи предполага определяне на най-информативни думи от текста, заглавието и/или резюмето на документа, които характеризират многоаспектно неговото съдържание. Езикът на ключовите думи се използва от електронните библиотеки на университетите, на традиционните библиотеки и на БАН (виж ил. 3).

Илюстрация 3. Дигитална библиотека Университет за национално и световно стопанство.



Макар че търсенето в изследваните електронни библиотеки се провежда по данни от библиографското описание, в някои групи електронни библиотеки, то съдържа съвсем основни данни, най-често автор и заглавие (някои електронни библиотеки на физически лица и на обществени организации, виж например Електронна библиотека „Моята библиотека“ ил. 4).

Илюстрация 4. Библиографско описание от Електронна библиотека „Моята библиотека”.



Библиографски описания, които отговарят на приети стандарти, съставят електронните библиотеки на обществените библиотеки и университетите. В някои библиотеки описанието може да бъде представено във формат ISBD, (например виж Дигитална библиотека „Мост през вековете”– Търговище, ил. 5).

Илюстрация 5. Формат за показване ISBD в Дигитална библиотека „Мост през вековете”– Търговище.



The image shows a digital library interface. On the left is a thumbnail of a newspaper page titled "ГРАЖДАНСКО СЪЗНАНИЕ" with a sub-headline "ВЪ БЪЛГАРИЯ". On the right is a metadata panel. At the top, a red header reads "Информация за броя:". Below it, a list of details includes: "тип: вестник", "заглавие: Гражданско съзнание", "подзаглавни данни:", and "дата: 15 дек. 1929, No 15". A tab labeled "ISBD" is selected. Below this, another red header reads "Информация за периодично издание:". The main text describes the publication as "ГРАЖДАНСКО СЪЗНАНИЕ : Безплатен народен вестник за култура и общинска политика : Излиза периодично / Ред. Н[еделчо] Куюмджиев, Д. Свинаров, А[сен] Каваев. - Год. 1, бр. 1 (1929) - Год. 2, бр. 28 (1931). - Ески Джумая : п-ца Кареев, 1929-1931. - 41 см." At the bottom, a table lists two volumes: "Год. 1 1929 - 1930 3 февр. - 12 септ. 1 - 52" and "Год. 2 1930 - 1931 23 септ. - 25 юни 1 - 28".

Обобщени данни за използваните информационни езици в изследваните електронни библиотеки са представени в табл. 12.

Таблица 12. Информационни езици в българските електронни библиотеки.

Информационни езици	Групи електронни библиотеки – ЕБ (брой)					
	ЕБ на физически лица (4)	ЕБ на организации (4)	ЕБ на инст. БАН (4)	ЕБ на университети (5)	ЕБ на общински обществени библиотеки (7)	Общо
Език на библиографско описание	4	4	4	5	7	24
Език на йерархичните класификации	-	-	-	2	-	2
Език на предметните рубрики	1	2	-	1	3	7

Дескрипторен език	-	-	-	-	-	0
Език на ключовите думи	-	-	2	3	2	7

Анализът на данните от таблицата очертава няколко важни извода:

1. Първият е свързан с честотата (или активността) на използване на лингвистичните средства:

- всички изследвани електронни библиотеки използват езика на библиографските данни, който е натоварен с различни функции – идентифициране, структуриране, търсене, преобразуване;

- често за търсене и представяне семантиката на документите се използва езикът на ключови думите – 7 библиотеки и езикът на предметните рубрика – 7 библиотеки;

- рядко се използват традиционните библиотечни класификации за систематизация – 2 библиотеки използват УДК;

- не се използва дескрипторен език.

Считаме, че основните причини за това се коренят в спецификата на тези езици: изискват се специално обучени индексатори; индексването не може да бъде автоматизирано напълно (УДК); необходимо е разработване и поддържане на нормативни речници (класификация, тезаурус).

2. Резултатите от изследването на информационните езици в българските електронни библиотеки представени по групи библиотеки позволи да се очертаят следните тенденции:

- електронните библиотеки, създадени от физически лица и обществени организации и учреждения се характеризират с опростеност на лингвистичното осигуряване и слаба систематизираност на информационния масив. Обикновено търсенето е по думи от библиографското описание, най-често автор и заглавие;

- електронните библиотеки, създадени от традиционни библиотеки се отличават с богат апарат за търсене. Използват се различни информационни езици. Наред с широко приетото търсене по думи се използват и контролирани езици, които осигуряват значително по-ефективно търсене. Това важи и за електронните библиотеки на университетите.

Заклучение от анализа на данните от изследването

Анализът от резултатите от изследването на лингвистичното осигуряване на българските електронни библиотеки показва следните негативни тенденции:

- за представяне характеристиките на документите с цел идентифициране и търсене, не се използват приети стандарти;
- не се използва комплекс от информационни езици за осигуряване на различни видове търсене – библиографско търсене и тематично търсене. Обикновено се използва само един език за отразяване съдържанието на документите;
- в малка степен се използват кодовите класификационни езици, които имат предимството да структурират и систематизират информационния масив по удобен начин;
- не се използват възможностите на езиците за структуриране на документалния масив по тема (предмет, дисциплина, ключови думи).

Това води до неефективно търсене (информационен шум, непълнота и неточност на резултатите), а с увеличаване на обема на колекциите (информационния масив), този проблем ще се задълбочава. Решението е включване в лингвистичното осигуряване на електронната библиотека на изпитаните и наложените в библиотечната практика езикови средства – предметна класификация, йерархична класификационна схема по области на знанието, дескрипторен език, които предполагат използването на контролиран речник. Те биха осигурили удобна структура на информационния масив и биха повишили резултативността на търсенето.

3.5. Модел на лингвистично осигуряване на българска електронна библиотека с универсална тематика

Моделът на лингвистично осигуряване на българска електронна библиотека с универсална тематика е конструиран въз основа на:

- 1) Теоретичния анализ на информационните езици като специални изкуствени езици, създадени и използвани за представяне на характеристиката на документите и , разкриване на тяхното съдържание с цел идентифициране, търсене и възможност за удобно навигиране;
- 2) Определяне на ролята на отделните езици в лингвистичното осигуряване на електронна библиотека: цялостно представяне на характеристиките на документите,

представяне на предмета (темата) на документите и многоаспектно представяне на съдържанието на документите.

3) Извеждане на предимствата и ограниченията на информационните езици (Приложение 3);

4) Анализа на резултатите от изследването на лингвистичното осигуряване на българските електронни библиотеки в българското интернет пространство.

Прегледът на различни информационни източници, в т.ч. най-добрите практики за реализиране на електронни библиотеки, запознаването с теорията на информационните езици и анализът на използваните информационни езици в българските електронни библиотеки позволяват да предложим модел на лингвистично осигуряване, приложим за българска електронна библиотека. Този модел може да подпомогне разработчиците на електронни библиотеки в България. Моделът включва конкретни „стандарти”, признати и наложени в международната и националната библиотечна практика и представени на български език. Под „стандарт”, когато става въпрос за информационни езици, имаме предвид конкретен инструмент, който се използва широко (общоприет) в определено съобщество²⁵⁰. Моделът е приспособим за различни електронни библиотеки (без ограничение на език, тематика, тип на медията и вид на документите и др.), като могат да бъдат добавени или заменени някои стандарти с оглед спецификата на конкретната библиотека.

При конструирането на модела сме се водили от някои важни принципи, които гарантират неговата практичност и възможност за многократната приложимост:

1) Отчитане на международната практика при избора на конкретни информационни езици (инструменти), което гарантира международен обмен на данни и интегриране на различни метаданни и колекции от документи.

2) Съобразяване с националната практика при избора на информационни езици за лингвистичното осигуряване на електронната библиотека. Някои инструменти за отразяване на съдържанието на документите са свързани тясно с естествен език и национална култура.

3) Придържане към вече съществуващи стандартни информационни езици.

4) Комплексно използване на различни по тип/вид и предназначение информационни езици за отразяване на съдържанието на документите, за да се осигурят максимални възможности за търсене:

²⁵⁰ Broughton, Vanda. *Essential thesaurus construction*. London: FacetPublishing, 2006, p. 8.

– комплексно използване на *неконтролирани и контролирани езици*, съобразно техните предимства и недостатъци. Първите отразяват съдържанието на документите с думи от текста, заглавието и/или резюмето на документа. Те се характеризират с простота на използване, скорост на индексирание, гъвкавост при търсене, но имат някои ограничения, като ниска степен на точност при търсене, липса на контекстуалност (неясни отношения между термините), неудобни за навигиране (липса на йерархия) и др. Контролираните езици индексират документите с термини от речник и прилагат синтактични правила за свързване на термини. Това предполага по-точни и пълни резултати при търсене. Контролираните езици предлагат ценна помощ на потребителите при търсене и удобство при навигиране. Недостатъците на тези езици са свързани с тяхната сложност, скъпа и трудоемка поддръжка на речник (виж Приложение 3);

– комплексно използване на *кодови йерархични класификационни езици и словесни езици*. Кодовите класификационни езици осигуряват систематично представяне на документалния масив с лесно и разбираемо придвижване в йерархията (по-обща и по-тесни понятия). Те са много подходящи за организиране на навигиране по тема в електронната библиотека. Като основни ограничения на кодовите езици се сочат а) неизбежното изоставане от бързия темп на развитие на науката и терминологията, свързана с нея, поради сложността на йерархичната структура и б) процесът на индексирание на документите се автоматизира много трудно. Словесните (контролирани) езици предоставят възможност за азбучно подреждане на термините, за навигиране и направляват търсенето чрез семантичните връзки между понятията. Основен недостатък е загуба на информация при търсене на лексика, която не е включена в контролирания речник (Приложение 3);

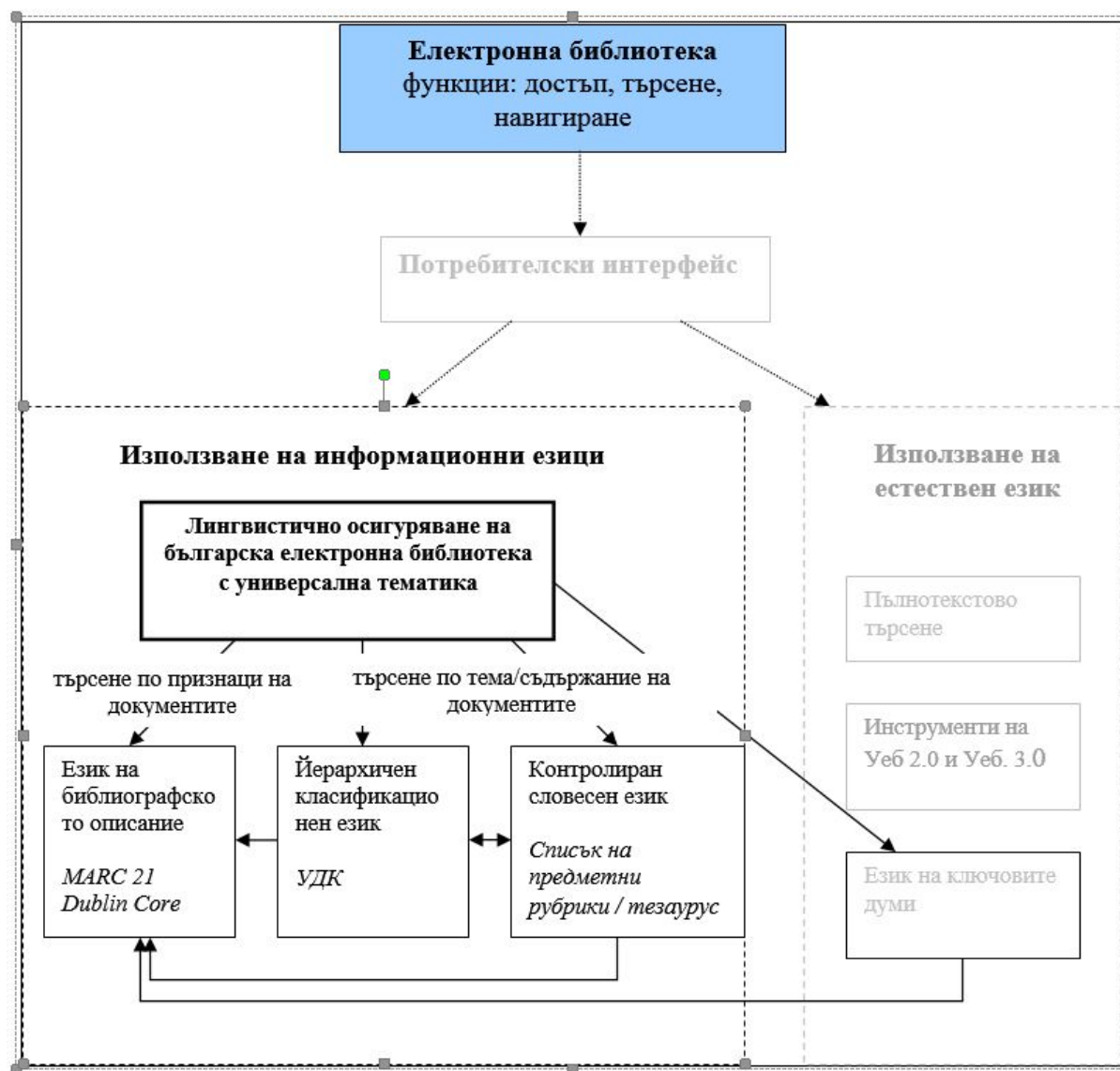
5) Изборът на конкретен език за описание (стандарт, схема, формат) може да зависи от избрания софтуерен продукт за създаване и поддръжане на електронна библиотека*. Затова в конструирания модел на лингвистично осигуряване ще разглеждаме само най-широко приложимите, които могат да бъдат адаптирани и използвани като обменни формати.

* Например политика на DSpace е да използва стандартни схеми за описание за постигане на съвместимост между DSpace и други системи и хранилища на метаданни с цел обединяването и експорт на данни. DSpace може да работи с MARC и Dublin Core виж: **Defining a Metadata Schema for the MIT Press Release of DSpace (DRAFT)** [online]. Viewed 20.06.2018]. Available from: <http://web.mit.edu/dspace-dev/www/Metadata-schema.htm>

б) Предвиждане на допълнителни методи и средства за търсене чрез използване на естествен език, които увеличават функционалността на електронната библиотека – пълнотекстово търсене, използване на технологии, подобряващи търсенето на Уеб 2.0 и Уеб 3.0.

Схемата на модела на лингвистично осигуряване на българска електронна библиотека с универсална тематика е представен на фиг. 21.

Фигура 21. Схема на модел на лингвистично осигуряване на българска електронна библиотека с универсална тематика



Моделът на лингвистично осигуряване на българска електронна библиотека с универсална тематика включва различни информационни езици с оглед тяхната специфична роля и осигуряване на максимални възможности за търсене: език на

библиографското описание; йерархичен класификационен език; контролиран словесен език; език на ключовите думи.

Изборът на всеки отделен език в модела се базира на използването му в международната практика.

Език на библиографското описание

Езикът на библиографското описание служи за обща характеристика на документите с цел тяхното идентифициране и търсене. Библиографското описание съдържа индекси на информационни езици, които отразяват съдържанието на документа – предметни рубрики, дескриптори, ключови думи, класификационни индекси. Затова връзките от тези езици към езика на библиографското описание са отразени в схемата на лингвистичното осигуряване като еднопосочни връзки (виж фиг. 23)

Международна практика

1. *MARC формати*. За представяне на библиографска информация и за обмен на данни библиотеките вече няколко десетилетия използват MARC формати (MARC 21 и UNIMARC и базирани на тях национални формати). Кодирането на MARC формат с език за маркиране XML го прави приложим за електронни библиотеки. Като обменни формати могат да бъдат използвани MarcXchange и MARCXML.

MarcXchange е международен стандарт, преведен и приет като национален стандарт и в България²⁵¹ и предназначен за:

- представяне на библиографско описание от формат MARC в XML;
- обмен на описания във формат MARC в XML;
- съставяне на оригинални описания на документи в синтаксис XML;
- трансфер на записи във формат MARC в уеб услуги като SRU (search/retrieval via URL);
- като временен формат при всякакъв вид трансформиране или управление на данни, напр. конверсия, публикуване, редактиране, валидиране;

²⁵¹ БДС ISO 25577:2015 Информация и документация. MarcXchange . В: *Български институт по стандартизация* [онлайн]. [Прегледан 20.06.2018]. Достъпно от: http://www.bds-bg.org/bg/standard/?natstandard_document_id=74166

– метаданни в XML, които могат да бъдат прикачени към електронен ресурс. MARCXML²⁵² служи и за конвертиране на данните и тяхното разполагане от MARC в друга описателна схема, например Dublin Core²⁵³.

2. *Dublin Core Metadata Standard* е първият стандартизиран формат за описание на електронни документи. Неговото начало е поставено през 1995 г. в Дъблин, щата Охайо и е известен като Множество от елементи метаданни Dublin Core. Основната цел за създаването му е изграждане на международен стандарт, който да подпомага търсенето и разпространението на информация. Той описва характеристиките на ресурсите, посочва тяхната интелектуална собственост и представя съдържанието им. Може да се комбинира с различни езици за маркиране като XML или RDF²⁵⁴.

3. *MODS (Metadata object description schema)*. „Описателната схема за метаданни на обекти” е разработена и поддържана от Библиотеката на Конгреса²⁵⁵. Тя позволява създаване на описателни метаданни. С нейна помощ се извличат данни от записи във формат MARC 21, а също така се създават и оригинални записи. Тази схема се използва в комбинация с език за маркиране XML. Елементите включени в нея са по-богати в сравнение със схемата Dublin Core, съдържайки и елементи за класификационен индекс, резюме, съдържание, предназначение и др.

4. *METS (Metadata encoding and transmission standard)*. Стандартът е предназначен за кодиране на описателни, административни и структурни метаданни за обекти в език за маркиране XML. Той се поддържа, като част от форматите MARC, от Конгресната библиотека. Състои се от седем големи секции: заглавна част на METS; описателни данни; административни данни; секция на файла; карта на структурата; връзки в структурата; информация за действие²⁵⁶.

5. *MADS (Metadata authority description schema)*. Описателна схема за нормативни метаданни представлява XML версия на „MARC21 Authority format”²⁵⁷. Тя служи за създаване и обмен на контролни данни за имена на лица, организации, предметни

²⁵² *MARCXML: MARC 21 XML Schema* [online]. [Viewed 20.06.2018]. Available from: <http://www.loc.gov/standards/marcxml/>

²⁵³ **Яврукова**, Биляна. *Библиографски контрол в процеса на ретроспективна конверсия и дигитализация на фондовете: Дисертация*. София, 2015, с. 93.

²⁵⁴ **Койчева**, Елена. The Dublin Core Metadata Initiative стандартизиран формат за метаданни за интернет пространството. В: *Computer*, 2004, № 12, с. 56.

²⁵⁵ **Outline** of Elements and Attributes in MODS Version 3.5. In: *MODS: Metadata Object Description Schema* [online]. [Viewed 20.06.2018]. Available from: <http://www.loc.gov/standards/mods/mods-outline-3-5.html>

²⁵⁶ Пак там.

²⁵⁷ **MADS Schema & Documentation** In: *MODS: Metadata Object Description Schema* [online]. [Viewed 20.06.2018]. Available from: <http://www.loc.gov/standards/mads/>

рубрики и др. Създадена е и се поддържа от Библиотеката на Конгреса. Последната актуализация на схемата MADS Version 2.1 е от 2016 г.

Представените схеми за създаване на метаданни са само една малка част от съществуващите. Наличието на множество схеми води до липса на стандартизация и унификация на използваните стандарти. При сравнение на елементите, съдържащи се в някои схеми, се вижда, че повечето от тях покриват изискванията за идентифициране и търсене на документ и са напълно съвместими с утвърдени схеми за метаданни като форматите MARC.

Приложимост в модела на лингвистично осигуряване на българска електронна библиотека с универсална тематика

MARC форматите имат следните предимства:

- те са „стандарт“ в света на библиотеките, наложен с времето*;
- достатъчно са гъвкави, за да описват различни типове/видове документи;
- развити са с последните технологии на уеб – MARCXML;
- могат да се използват във всяко от избраните инструментални средства за създаване на електронна библиотека (виж Глава 1, т. 1.5).

Българските библиотеки имат натрупан опит в създаването на библиографски записи с машинночетими формати. Форматите MARC21 и UNIMARC са преведени на български език²⁵⁸. MARC21 се използва в електронните каталози на едни от най-големите библиотеки в България (Централна библиотека БАН** и Библиотеката на Софийския университет***).

От прегледа на световната практика стана ясно, че в голяма степен се използват форматите MARC за организиране на библиографската информация в електронните информационно-търсещи системи на библиотеките. Използването на MARC в уеб среда е гарантирано от специално създадени обменни формати за представяне в XML и за обмен на описателни метаданни MarcXchange. Формати от семейството на MARC се използват като национални в страни, водещи в развитието на информатизацията на библиотеките.

* използва се от края на 60-те години.

²⁵⁸ Преводи на MARC21: *MARC 21 Lite* Формат за машинночетими библиографски записи. София: Унив. изд. „Св. Климент Охридски“, 2005; *MARC 21* формат за библиографски данни. Актуализация № 1 (октомври 2001 г.) – актуализация № 13 (септември 2011 г.). Достъпно от: <http://www.nalis.bg/MARC-VIB/>; Превод на UNIMARC: *Ръководство за UNIMARC*. София: Съюз на библиотечните и информационни работници, 2002. Достъпно от: <http://www.lib.bg/prevodi/unimarc/>

** <http://cl.bas.bg/informacionni-uslugi/on-line-katalog>

*** http://aleph.libsu.uni-sofia.bg:8991/F/?func=find-b-0&con_lng=bul

Посочените предимства определят включването на MARC формат за описание на характеристиките на документите в модела на лингвистично осигуряване на българска електронна библиотека. Затова ще представим двата най-използвани в света машинночетими формати за библиографски записи и преведени на български език.

MARC21

„MARC21 Формат за библиографски данни” е най-разпространеният машинно четим формат за описание от семейството на MARC форматите²⁵⁹. Той е създаден с цел стандартно организиране на библиографската информация при каталогизиране на различни видове информационни ресурси и за обмен на библиографска информация. Форматът служи за съставяне на библиографски записи за седем типа документи: книги, продължаващи издания, файлове, географски карти, музика, визуални материали, смесени материали.

Според типа на каталогизирания документ MARC 21 определя 14 различни типа записи, съдържащи различен набор от елементи от библиографски данни, които трябва да се запишат в MARC записа: текстови документи, ръкописни текстове, нотни издания, ръкописни ноти, картографски издания, ръчно изработени карти, прожекционни материали, звукозаписи различни от музика, звукозаписи на музика, графични издания, файлове, комплекти, смесени материали, триизмерни обекти/артефакти²⁶⁰.

Библиографските данни, с които се описват документите, съдържат заглавия, имена на лица и названия на организации, названия, издателски данни и физическа данни, данни за съдържанието на документа (предметни рубрики, индекси по класификационна схема, ключови думи, дескриптори), връзки с други библиографски записи и други.

Библиографските данни в MARC 21 са групирани и представени в следните блокове от структурирана информация²⁶¹:

0XX Полета за контрол, Контролна информация и Кодови класификационни индекси

1XX Основни точки за достъп

2XX Заглавие, авторска отговорност и Издателски данни (обща информация)

²⁵⁹ Миланова, Милена. *Българската каталогизация в глобалното ...*, с. 64.

²⁶⁰ Лидер (NR). В: *MARC 21 формат за библиографски данни: Кратка версия* [online]. 2010. [Прегледан 20.06.2018]. Достъпно от: <http://www.nalis.bg/MARC-BIB/concise/bdleader.html>

²⁶¹ *MARC 21 Format for Bibliographic Data* [online]. 1999 Edition Update No. 1 (October 2000) through Update No. 26 (April 2018) Достъпно от: <http://www.loc.gov/marc/bibliographic/>; на български език: *MARC 21 формат за библиографски данни*. Актуализация № 1...

- 3XX Физическа характеристика и др.
- 4XX Серия
- 5XX Забележки
- 6XX Предметни рубрики
- 7XX Допълнителни точки за достъп (имена, названия) и Връзка между записите
- 8XX Точки за достъп за серии, служебни данни и местонахождения
- 9XX За национално използване

Съдържанието на документа в MARC21 се отразява:

а) чрез индекси на кодови класификации в блок 0XX, полета 050–084:

- 050 Класификация на Конгресна библиотека (Library of Congress Subject Headings)
- 055 Класификационна система на библиотеките и архивите в Канада (Classification systems of Library and Archives Canada)
- 070 Класификационна система на Националната селскостопанска библиотека на САЩ (Classification systems of National Agricultural Library)
- 086 Класификационна система на U.S. Government Printing Office (Superintendent of Documents classification system of U.S. Government Printing Office)
- 080 Универсална десетична класификация (Universal Decimal Classification)
- 082 Десетична класификация на Дюи (Dewey Decimal Classification schemes)
- 084 Други класификации, (индекси на класификации, които са регистрирани и им е присъден код)*.

б) чрез словесни информационни езици в блок 6XX:

– индекси на предметизационни схеми/дескриптори - полета 600–651 (600, 610, 611, 630, 647, 648, 651), в зависимост от съдържанието им (име на лице, наименование на организация, тема, географски обект и др.) ;

– ключови думи поле 653;

– индекси от фасетна класификация – поле 654.

„MARC21 Формат за библиографски данни” регламентира използването на няколко предметни класификации или тезауруси чрез отразяването им в индикатор на посочените полета (600–651):

Предметни рубрики на Конгресна библиотека (Library of Congress Subject Headings)

Предметни рубрики на Конгресна библиотека за детска литература (LC subject headings for children's literature)

Предметни медицински рубрики (Medical Subject Headings)

* *вж Classification Scheme Source Codes, <http://www.loc.gov/standards/sourcelist/classification.html>*

Предметни рубрики на Националната селскостопанска библиотека на САЩ (National Agricultural Library subject authority file)

Канадски предметни рубрики (Canadian Subject Headings)

Предметни рубрики на Библиотека и архив на Канада Répertoire de vedettes-matière.

Други предметни класификации, получили код от Конгресната библиотека, могат да се използват в подполе \$2 на съответните полета*.

За да бъде машинночетима класификационната схема е създаден специален формат „MARC21 Формат за класификационни данни” (*MARC21 Format for Classification Data*)²⁶². Класификационните данни в MARC формат могат да се използват за търсене, за автоматизиране на процеса на индексване, за създаване на класификационни схеми, за валидиране на присвоените класификационни кодове в библиографски запис²⁶³. „MARC21 Формат за класификационни данни” е разработен да отговаря на двете основни класификационни схеми, използвани в САЩ и значителна част от англоговорящите страни: Десетична класификация на Дюи и Класификация на Конгресна библиотека, но е приложим и за други класификации²⁶⁴.

За контролираните речници на предметните класификации и тезаурусите се използва „MARC 21 Формат за нормативни данни (*MARC21 Format for Authority Data*). Той служи за структурирано представяне на нормативна информация – форми на имена (лица, организации, временни колективи, заглавия) и на предметни рубрики (теми, географски понятия, имена на лица, наименования на организации и т.н.). „MARC 21 Формат за нормативни данни” е създаден за най-използваните от американските библиотеки предметизационни схеми: предметните рубрики на Библиотеката на Конгреса на САЩ (*Library of Congress Subject Headings*), списъкът с предметни рубрики на Сиърс (*Sears List of Subject Headings*), тезаурусът за изкуство и архитектура (*Art and Architecture Thesaurus*) и др., но е приложим за машинночетимото представяне различни предметизационни схеми и тезауруси.

* виж *Subject Heading and Term Source Codes*, <http://www.loc.gov/standards/sourcelist/subject.html>

²⁶² <https://www.loc.gov/marc/classification/>

²⁶³ <https://www.loc.gov/marc/classification/cdintro.html>

²⁶⁴ <http://www.nalis.bg/MARC-UA/pt1-7.html#table1>

Dublin Core

Dublin Core Metadata Element Set (DCMES)* е приет за стандарт от различни национални и международни организации: IETF RFC 5013, стандарт ISO 15836-2009, NISO стандарт Z39.85. В България функционира като БДС ISO 15836-1:2017 Информация и документация. Множество от елементи на метаданни DUBLIN CORE. Част 1: Основни елементи.

Dublin Core представлява кратък набор от елементи от библиографски данни за описание на разнообразни информационни обекти, представени в следната типология (DCMI Type Vocabulary)**:

Колекция (Collection) – съвкупност от ресурси, описани общо. Отделните единици в колекцията могат да бъдат описани самостоятелно/отделно.

Набор от данни (Dataset) – данни, представени в определена структура например списъци, таблици или бази от данни. Наборът от данни може да бъде предназначен за директната обработка от машината.

Събитие (Event) – случка, явление, обвързано с време. Примерите включват изложба, конференция, семинар, ден на отворените врати, представление, битка, процес, сватба, пожар и т.н. Метаданните за „събитие” предоставят описателна информация за цел, местоположение, продължителност, обвързани със събитието лица и др.

Изображение (Image) – ресурс представен визуално, примерите включват вдижещи се и статични изображения – снимки на физически обекти, картини, отпечатащи, рисунки, графики, диаграми, карти, музикални ноти, анимации, филми и др.

Движещо се изображение (MovingImage) – поредица от изображения, които придават впечатление за движение, когато се показват последователно. Примерите включват анимации, филми, телевизионни програми, видеоклипове, визуални симулации.

Статично изображение (StillImage) – неподвижно визуално представяне например картина, рисунка, графичен дизайн, планове и карти.

Звук (Sound) – ресурс, предназначен за слушане например формат за възпроизвеждане на музика, компактен аудио диск, запис на говор или звуци.

* <http://dublincore.org/documents/dces/>

** <http://dublincore.org/documents/dcmi-type-vocabulary/>

Текст (Text) – ресурс, състоящ се основно от думи, символи, цифри за четене. Примерите включват книги, писма, дисертации, стихотворения, вестници, статии, архиви на пощенските списъци, факсимилита или изображенията на текстове.

Интерактивен ресурс (Interactive resource) – ресурс, който изисква взаимодействие от потребителя например уеб страници, аплети, мултимедийни учебни обекти, чат услуги или среда за виртуална реалност.

Физически обект (Physical object) – реален триизмерен предмет.

Услуга (Service) – една или повече функции, предоставяни от система например услуга за удостоверяване на автентичността, междубиблиотечно заемане и т.н.

Софтуер (Software) – компютърна програма.

Основният набор DCMES съдържа 15 основни елемента. Всеки елемент има описателен етикет за разпознаване от човек и уникален знак – наименование за употреба при машинна обработка. Наименованието на елемент е прикрепено към унифицирания идентификатор на ресурса (URI) в именното пространство – DCMI namespace²⁶⁵:

Заглавие – наименование, дадено на ресурса; Етикет: Title.

Създател – субект, носещ първоначалната отговорност за създаването на ресурса; Етикет: Creator

Значение: Примерите за „създател” включват лице, организация или служба.

Предмет – тема на ресурса ; Етикет: Subject

Значение: Обикновено, предметът се изразява като се употребяват ключови думи, предметни рубрики или класификационни кодове. Най-добрата практика е да се употребява контролиран речник.

Описание – изложение на съдържанието на ресурса ; Етикет: Description

Значение: Описанието може да включва, например реферат, съдържание и др. п, които представят съдържанието чрез свободен текст.

Издател – субект, отговорен за предоставяне на достъп до ресурса ; Етикет: Publisher

Значение: Примерите за „Издател” включват лице, организация или служба.

Участник – субект, отговорен за даването на своя принос при създаването на ресурса ; Етикет: Contributor

Значение: Примерите за „Участник” включват лице, организация или служба.

²⁶⁵ Съгласно RFC 3986 URI Generic Syntax, достъпно от: <https://www.ietf.org/rfc/rfc3986.txt>

Дата – определен момент или период от време, свързани с дадено събитие в жизнения цикъл на ресурса ; Етикет: Date

Значение: Добра практика за изразяване на датата, времето или периода е съобразно стандарт „Елементи от данни и формати за обмен. Информационен обмен. Представяне на дати и време БДС ISO 8601:2006”.

Тип – характерът или жанрът на ресурса ; Етикет: Type.

Значение: Добра практика е използването на контролиран речник като DCMI Type Vocabulary.

Формат – формат на файла, физически носител или величините на ресурса ; Етикет: Format.

Значение: Примерите за величини включват големина и времетраене. Добра практика е използването на контролиран речник, напр. Internet Media Types (MIME).

Идентификатор – кратка последователност от писмени знаци, която създава идентичността на ресурса ; Етикет: Identifier.

Значение: Добра практика е да се идентифицира ресурсът чрез низ, съобразен с официална система за идентификация.

Източник – сроден ресурс, от който произхожда описваният ресурс ; Етикет: Source

Значение: Описваният ресурс може да произхожда изцяло или частично от сродния ресурс. Най-добрата практика е да се идентифицира сродния ресурс чрез низ, съобразен с официална система за идентификация.

Език – език на ресурса ; Етикет: Language.

Значение: Добра практика е да се употребява контролиран речник, напр. БДС ISO 639-1:2017 Кодове за представяне на наименованията на езиците. Част 1: Код алфа-2.

Връзка – свързаност със сроден ресурс ; Етикет: Relation.

Значение Добра практика е да се идентифицира сродният ресурс чрез даден URI (унифицирания идентификатор на ресурса) или друг низ, съобразен с официална система за идентификация.

Обхват – пространство, време или юрисдикция, свързани с темата на ресурса ; Етикет: Coverage.

Значение: Добра практика е да се употребява контролиран речник, напр. Тезаурус на географските наименования - Getty Thesaurus of Geographic Names (TGN).

Права – информация за правата по ограничаване на достъпа и защита на ресурса ; Етикет: Rights.

Значение: Информацията за правата обикновено включва изразяване на различни права на собственост, свързани с ресурса, включително и правата на интелектуалната собственост.

Простотата на набора от елементи се сочи като основна причина за широкото разпространение и приложение на Dublin Core²⁶⁶. Същевременно „минималността“ и неопределеността на елементите се определят като съществен недостатък²⁶⁷. Компромисно решение е използването на Dublin Core определители (Qualifiers), които се използват към набора от елементи. Те са два вида: спецификация (Element Refinement) и схема (Encoding Scheme). Първият вид определители е предназначен да специфицира значението на елементите, а вторият да посочват схема или контролен списък, който се използва при представяне стойността на елемент (табл. 13).

Таблица 13. DCMES елементи и определители

DCMES Елемент	Спецификация	Схема / Контролиран речник
Заглавие	Алтернативно заглавие	-
Създател	-	-
Предмет	-	Предметни рубрики на Конгресна библиотека (LCSH) Медицински предметни рубрики (MeSH) Десетична класификация на Дюи (DDC) Класификация на Конгресна библиотека (LCC) Универсална десетична класификация (УДК)
Описание	Съдържание Резюме	-
Издател	-	-

²⁶⁶ Миланова, Милена. *Българската каталогизация в глобалното ...*, с. 97.

²⁶⁷ Rust, Godfrey. Metadata: The Right Approach. In: *D-Lib Magazine* [online], July/August 1998. [Viewed 21.06.2018]. Available from: <http://www.dlib.org/dlib/july98/rust/07rust.html>

Участник	-	-
Дата	Дата на създаване Дата на валидиране Дата на достъп Дата на издаване Дата на модифициране	DCMI Period БДС ISO 8601:2006 Представяне на дати и време W3C Date and Time Formats
Тип	-	DCMI Type Vocabulary
Формат	Размер	-
	Медия	Internet media type (MIME)
Идентификатор	-	URI
Източник	-	URI
Език	-	БДС ISO 639-2 Кодове за представяне на наименованията на езиците. Част 2: Код Алфа-3 RFC 1766 Етикети за идентификация на езици
Връзка	Е версия на Има версия Е заменен от Заменя Изискван от Изисква Е част от Има част Е цитиран от Цитира Е формат на Има формат	URI
Обхват	Пространствен	DCMI Point БДС ISO 3166 Кодове за представяне на наименованията на държавите и техните подразделения DCMI Box

		Тезаурус на географските наименования (TGN).
	Времеви	DCMI Period W3C Date and Time Formats
Права	-	-

За отразяване на съдържанието на документите и тематично търсене в Dublin Core се използва елемент Предмет. Стандартът препоръчва предмета (темата) да се изрази чрез стандартни информационни езици (кодови и словесни) с контролиран речник – Предметни рубрики на Конгресна библиотека, Медицински предметни рубрики, Десетична класификация на Дюи, Класификация на Конгресна библиотека, Универсална десетична класификация.

Стандартът Dublin Core може да се прилага за: библиографско описание на ресурсите и за търсене; за комбиниране на речници на метаданни на различни стандарти за метаданни, осигуряване на оперативна съвместимост. Той е напълно съвместим с MARC форматите, което прави възможно конвертирането на данните от един формат към друг. Съпоставка на полетата на MARC 21 и елементите на Dublin Core с определителите е представена в Приложение 4. Представените характеристики на схемата за метаданни Dublin Core са достатъчно основание за нейното включване в модела на лингвистично осигуряване на българска електронна библиотека

Изборът на конкретен език за описание (стандарт, схема, формат) може да зависи от избрания софтуерен продукт за създаване и поддържане на електронна библиотека*. Затова тук разглеждаме само най-широко приложимите, които могат да бъдат адаптирани и използвани като обменни формати.

Йерархичен класификационен език

Йерархичните класификации представляват универсален инструмент за организация на знанието с удобна цифрова система за обозначаване на термините с ясно

* например политика на DSpace е да използва стандартни схеми за описание за постигане на съвместимост между DSpace и други системи и хранилища на метаданни с цел обединяването и експорт на данни. На различни видове колекции от данни и да се позволи глобално търсене на повече от една система през шлюзове и протоколи. DSpace може да работи с MARC и Dublin Core виж: *Defining a Metadata Schema for the MIT Press Release of DSpace (DRAFT. Purposes and Types of Metadata* <http://web.mit.edu/dspace-dev/www/Metadata-schema.htm>

изразени йерархични и смислови връзки между термините. Те имат важна роля за организирането на информационния масив по области на знанието и дисциплини в електронната библиотека. Те предполагат организиране чрез кодовата нотация и търсене чрез словесния израз на индекса. Електронната библиотека може да използва само най-високите йерархични нива за структуриране на информационния масив. Възможно е и да се индексират документите с пълен класификационен индекс, който се отразява в библиографското описание. Той служи за по-детайлно търсене чрез цифровата нотация или словесния израз. Главното предимство за използването на йерархична класификация в електронна библиотека е, че тя осигурява систематично представяне на тематичния обхват на библиотеката и предлага логично представяне на по-общи или по-конкретни теми като лесно може да се види връзката между тях.

В международен план класификационните схеми осигуряват потенциален многоезичен достъп²⁶⁸. Цифровите или буквено-цифровите обозначения на понятията са разбираеми и не зависят от естествен език.

Използваните в библиотеките класификационни схеми се различават по обем и методология: могат да бъдат международни, национални и локални; да обхващат всички области на знанието или да бъдат предметно ориентирани. Препоръчително е използването на общи международни стандарти за тематичен достъп и търсене в библиотеките, което осигурява международен обмен на данни и интегриране на различни метаданни и колекции.

В случая с електронна библиотека с универсална тематика е необходима обща класификация, която да обхваща множество понятия от всички отрасли на знанието и човешката дейност. Ако се налага използването на класификация за отделни специфични колекции, то те също трябва да бъдат съгласувани с общоприетите класификации. Това се отнася и за класификациите създадени за национално използване. Решението за използването на конкретна йерархична класификация зависи от предметния обхват и обем на електронната библиотека и необходимата степен на детайлизация при индексирание.

²⁶⁸ IFLA. *Guidelines for Subject Access in National Bibliographies* [online]. Draft May 2011 for Worldwide Review. IFLA Working Group on Guidelines for Subject Access by National Bibliographic Agencies [Viewed 21.06.2018]. PDF, 7.5 MB. Available from: https://www.ifla.org/files/assets/classification-and-indexing/subject-access-by-national-bibliographic-agencies/nba_guidelines_draft_2011-05.pdf

Международна практика

В световен план най-използваните универсални йерархични класификации в библиографските агенции²⁶⁹ и в библиотеките, както традиционни така и електронни са Десетичната класификация на Дюи, Класификацията на Конгресна библиотека, Универсалната десетична класификация²⁷⁰. Специално трябва да бъде подчертано, че не съществува общ стандарт за тематично индексирание, разбиран като „публикувани нормативни критерии за качество на продукт или начин, който трябва да се следва при осъществяването на процес“²⁷¹. Широко разпространените международни класификации се приемат като общи стандартни средства.

Десетична класификация на Дюи (ДКД)*

Обща характеристика. ДКД е създадена от М. Дюи през 1873 г. и от нейното първо публикуване през 1876 г. до днес тя непрекъснато се актуализира, редактира и преиздава.

ДКД се издава в пълно и съкратено издание, както и в електронна версия от OCLC (Online Computer Library Center, Inc.). Съкратеното издание е логично съкращение на пълната таблица и е подходящо за универсални колекции с обем до 20 хил. документи. Класификацията е преведена на повече от тридесет езика, но не е превеждана на български език.

Достъпност. Информация за изданията и различните версии, за използването и приложението на ДКД е предоставена на сайта на OCLC** и в специализирания блог***. Класификационната схема е достъпна онлайн за регистрирани потребители WebDewey****. От 2018 г. OCLC предоставя експериментална уеб услуга в помощ на

²⁶⁹ IFLA. *Guidelines for Subject Access in National Bibliographies...*, p. 42.

https://www.ifla.org/files/assets/classification-and-indexing/subject-access-by-national-bibliographic-agencies/nba_guidelines_draft_2011-05.pdf

²⁷⁰ Shiri, Ali, Chase-Kruszewski, Sarah. Knowledge organisation systems in North American digital library collections. Abstract. In: *Program*, 2009, Vol. 43, Issue 2, pp.121–139. [Viewed 21.06.2018]. Available from: <https://www.emeraldinsight.com/doi/pdfplus/10.1108/00330330910954352>

²⁷¹ IFLA. *Guidelines for Subject Access in National Bibliographies...*, p. 42.

https://www.ifla.org/files/assets/classification-and-indexing/subject-access-by-national-bibliographic-agencies/nba_guidelines_draft_2011-05.pdf

* Dewey Decimal Classification (DDC)

** <https://www.oclc.org/en/dewey.html>

*** <http://ddc.typepad.com/>

**** <http://dewey.org/webdewey/login/login.html>

класификацията Classify*, която дава възможност за търсене по стандартен номер, автор, заглавие и предметна рубрика и предлага класификационен индекс по ДКД и по Класификация на Конгресна библиотека като показва най-използваните индекси за класифициране на търсения документ или предметна рубрика.

Ползватели. Над 200 000 библиотеки от всички типове от 135 страни използват ДКД за тематична организация и достъп до своите колекции²⁷².

Универсалната десетична класификация (УДК)

УДК е класификационна схема от йерархичен тип с цифрова нотация считана за международен стандарт в широкия смисъл, който В. Броутън влага в понятието „стандарт”, т.е. общоприетост в дадено съобщество²⁷³.

Обща характеристика. УДК е създадена от П. Отле и А. Лафонтен в края на XIX в. За основа на класификацията е използвана десетичната схема на Дюи. Новосъздадената класификация има усъвършенствана структура и по-голяма детайлизираност. Основната разлика между УДК и класификацията на Дюи е въвеждането на общи и специални определители и граматични средства, които позволяват да се създават нови понятия и да се изразят чрез индекси²⁷⁴. Отговорна организация за поддържането и развитието на УДК от момента на нейното създаване до 1980 г. е Международната федерация по информация и документация (Federation Internationale de Documentation, FID). През 1991 г. се създава независима международна организация отговорна за контрола по използването и издаването, както и за поддържането и развитието на класификацията. В Консорциума членове учредители са едни от най-активните издатели на национални версии на УДК: AENOR (Испания), BSI (Великобритания), CEFAL (Белгия), ВИНТИ (Русия) и Националната библиотека на Чехия²⁷⁵. Консорциумът поддържа еталонна база данни на таблиците на УДК на английски език (Master Reference File, MRF), която съдържа над 70 000 класове и около

* <http://classify.oclc.org/classify2/ClassifyDemo?wi=1006748964>

²⁷² **Break** down language barriers. In: *OCLC* [online]. Dewey services. Overview. [Viewed 21.06.2018]. Available from: <https://www.oclc.org/en/dewey.html>

²⁷³ **Broughton**, Vanda. *Essential thesaurus construction*. London: Facet Publishing, 2006, p 8.

²⁷⁴ *Универсална десетична класификация...*

²⁷⁵ **IFLA**. *Guidelines for Subject Access in National Bibliographies...*, p. 25.

250 хил понятия²⁷⁶. Еталонната база се актуализира и усъвършенства непрекъснато едновременно развитието на знанията и информационната дейност.

Достъпност. Информация за УДК, за нейните издания, версии и изменения, включително библиографски списък на публикации, посветени на изследването на класификацията, може да се намери на сайта на консорциума*. Онлайн услугата „UDC Online” предлага платен достъп до еталонната база данни, освен това включва 11 000 индекси, които са били променени или отменени и тяхното пренасочване към други индекси**.

Ползватели. УДК се използва от библиографски агенции, информационни и документални центрове и библиотеки от 130 страни в света. Тя е преведена на 40 езика²⁷⁷, включително на български език. Според проучване на Консорциума около 140 хил. библиотеки само в Европа използват УДК²⁷⁸.

Класификация на Конгресна библиотека (ККБ)***

*Обща характеристика*²⁷⁹. Историята на Класификацията на Конгресна библиотека е свързана с организацията на личната колекция от книги на Томас Джеферсън, която поставя началото на Конгресна библиотека. Неговата схема разделя знанието на три основни класове: история, философия и изкуство. Схемата е приета от библиотеката и развивана до края на XIX в. През 1901 г. експерти преработват основно класификацията, за да я адаптират към бързо растящия фонд, енциклопедичен по характер. В съвременния си вид ККБ се състои от 21 основни класове и 41 специални класове, обозначени с букви. В периода 1991–1996 г. класификацията е конвертирана в MARC формат.

Достъпност. Информация за ККБ може да се намира на сайта на Конгресна библиотека****. Класификацията е достъпна като онлайн услуга „Classification Web” в

²⁷⁶ UDC Master Reference File (MRF) <http://www.udcc.org/index.php/site/page?view=mrf>

* <http://www.udcc.org/index.php>

** <http://www.udc-hub.com/en/login.php>. Кратката онлайн версия на УДК е достъпна свободно на адрес <http://www.udccsummary.info/php/index.php>.

²⁷⁷ **Who** is using UDC? In: *UDC concertium* [online]. About Universal Decimal Classification (UDC). [Viewed 21.06.2018]. Available from: <http://www.udcc.org/index.php/site/page?view=about>

²⁷⁸ **UDC** Users Worldwide. In: *UDC concertium* [online]. [Viewed 21.06.2018]. Available from: http://www.udcc.org/index.php/site/page?view=users_worldwide

*** Library of Congress Classification (LCC)

²⁷⁹ **Mai**, Chan Lois. Library of Congress Classification in a New Setting: Beyond Shelfmarks. In: *Library of Congress* [online]. Classification Distribution Service. [Viewed 21.06.2018]. Available from:

<http://2008.myvote.org/www.loc.gov/cds/chanarticle.html>

**** <https://www.loc.gov/cds/classweb/>

помощ на класификаторите - може да се търси и навигира по буквено-цифровите индекси на класификацията, по словесните им изрази и по предметните рубрики на Конгресна библиотека.

Ползватели. Класификацията се използва от големи научни и университетски библиотеки в САЩ и Канада, Уругвай и Венецуела.

Редица страни създават собствени класификационни схеми, които отразяват националната им култура и език. Сред тях са Индия, Китай, Дания, Япония.

Приложимост в модела на лингвистично осигуряване на българска електронна библиотека с универсална тематика

В българската библиотечна практика УДК е позната още от 1903 г.²⁸⁰ и днес е широко разпространена като единна класификационна система²⁸¹. Това предопределя избора на УДК в състава на лингвистичното осигуряване на българска електронна библиотека с универсално съдържание.

УДК се характеризира с няколко основни свойства и принципи²⁸²:

Универсалност

Включени са множество понятия от всички отрасли на знанието и човешката дейност, представени по тематичен принцип в следните основни отдели:

0 НАУКА И ЗНАНИЕ. ОРГАНИЗАЦИЯ НА УМСТВЕНИЯ ТРУД. КОМПЮТЪРНИ НАУКИ. ИНФОРМАЦИЯ. ДОКУМЕНТАЦИЯ. БИБЛИОТЕЧНО ДЕЛО. ИНСТИТУЦИИ. ИЗДАНИЯ

- 1 ФИЛОСОФИЯ. ПСИХОЛОГИЯ
- 2 РЕЛИГИЯ. ТЕОЛОГИЯ
- 3 ОБЩЕСТВЕНИ НАУКИ
- 4 [свободен за бъдещо използване]
- 5 МАТЕМАТИКА. ЕСТЕСТВЕНИ НАУКИ
- 6 ПРИЛОЖНИ НАУКИ. МЕДИЦИНА. ТЕХНИКА
- 7 ИЗКУСТВО. РАЗВЛЕЧЕНИЯ. ЗАБАВЛЕНИЯ. СПОРТ
- 8 ЕЗИК. ЕЗИКОЗНАНИЕ. ЛИТЕРАТУРА
- 9 ГЕОГРАФИЯ. БИОГРАФИИ. ИСТОРИЯ.

²⁸⁰ Младенова, Мария. Цит. съч., с. 135.

²⁸¹ *Универсална десетична класификация...*, с. 8.

²⁸² Представени по: *Универсална десетична класификация...*

Макар отделите на класификацията, които съответстват на отделни отрасли на знанието, да са различни по своята специфична структура, системата се възприема като едно цяло заради единния йерархичен код, общите правила за изграждане на индексите и посочването на връзките на определен отдел с други отдели чрез методическия апарат. УДК е универсална и по отношение на прилагането си. Многобройните възможности и похвати за индексирание и за съкращаване на дробността позволяват тя да се използва успешно за систематизиране и последващо търсене на различни информационни ресурси в разнообразни по обем и предназначение фондове – от малки колекции от специализирана документация до големи отрасли или универсални библиотечни фондове.

Десетичност

Една от главните отличителни особености на УДК е йерархичната структура. В основата на класификацията е десетичният принцип на деление на цялото човешко знание.

Йерархичност

Йерархичният принцип на изграждане от общо към частно се отнася както за основната, така и за помощните таблици на общите и специални определители

Систематичност

УДК е изградена на систематичен принцип. Ако при предметната класификация всички сведения за предмета (понятието) са събрани на едно място, независимо от отрасъла на знанието, към което се отнасят, в УДК предметът може да се срещне на много места в класификационните таблици в зависимост от отрасъла на знанието и от аспекта, в който той е разгледан.

Многоаспектност

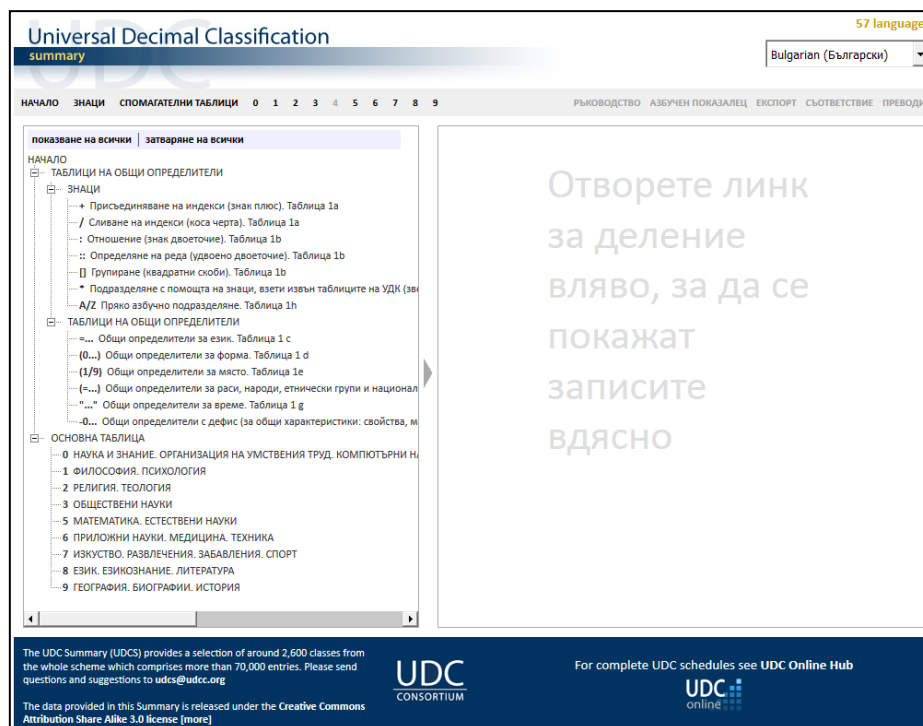
Множествената локализация на понятията отразява многоаспектността като свойство, заложено в самата структура на класификационната схема. Многоаспектното индексирание на съдържанието на документите и на информационните запитвания се осигурява от самата структура на УДК, от използването на общите и специалните определители, отразяващи време, място, език, параметри, оборудване, процеси и т.н., и от прилагането на някои правила за комбиниране на индекси.

Таблиците на УДК на български език са издавани няколкократно – през 1955, 1960, 1967 и 1985 г. През 1992 г. излиза преработка на отделите, структурно различни от

стандартното издание и нов Азбучно-предметен показалец на таблиците²⁸³. През 1996 г. е приет национален стандарт за подобрени деления на УДК (БДС 17388-96, от 2015 г. отменен). Последно излиза съкратено издание на таблиците на УДК през 2017 г.

Съкратена версия на УДК може да се използва на 57 езика (включително български) в онлайн режим в база от данни UDC Summary* (ил. 6. В изданието е направен подбор на около 2 600 деления, извлечени от пълното съдържание (Master Reference File) на УДК. Той включва основни индекси, общи и специални определители и представлява и покрива балансирано всички области на знанието, според степента на тяхното отразяване в класификационната система.

Илюстрация 6. UDC Summary.



UDC Summary е стандарт и официален съкратен вариант на Универсалната десетична класификация, който непрекъснато се актуализира с промените и допълненията, приемани в Master Reference File²⁸⁴. UDC Summary може пълноценно да се използва за индексирание на съдържанието на документите и информационните

²⁸³ Дипчикова, Александра, Тогоманова, Антоанета. *Промените в системите за организация на знанието...*

* <http://www.udcsummary.info/php/index.php?lang=bg#>

²⁸⁴ Пак там.

запитвания и за структуриране и систематизиране на информационните обекти в електронна библиотека.

Словесен контролиран език

Електронната библиотека трябва да разполага с ефективни инструменти за словесен тематичен достъп, съответстващ на международни принципи и стандарти²⁸⁵.

Индексирането на съдържанието на документите чрез информационен език с контролиран речник е предпоставка за провеждане на успешно търсене, тъй като речникът предоставя възможност да се направлява търсенето с по-широки или по-тесни термини, синоними, асоциативни термини. За тази цел се използва списък на предметни рубрики или тезаурус. Разликата между двата инструмента в тяхното развитие се заличава все повече.

Международна практика

Използването на международни широко разпространени предметизационни схеми съкращава усилията по индексирането и улеснява повторното използване на данни. Списъци с предметни рубрики и нормативни данни се създават и поддържат от националните библиотеки или националните библиографски агенции на страните. За разлика от йерархичните класификации, предметните рубрики и тезаурусите са свързани с естествен език и трябва да отразяват в достатъчна дълбочина теми с национален или местен аспект. Затова използването на международно разпространена предметизационна схема налага нейната адаптация, дори да не е нужен превод.

Най-популярната предметизационна схема е списъкът с предметни рубрики на Конгресната библиотека (Library of Congress Subject Headings, LCSH). Той се използва широко не само в САЩ, но и в целия свят, най-вече в англоезичните страни. Адаптации на LCSH са направили националните библиотеки на Австралия, Канада, Латвия, Литва и др.

Друг подход е съставянето на национална предметизационна схема или тезаурус, които да се използват в библиотеките на страната за стандартизиран и унифициран

²⁸⁵ Lucarelli, Anna, Cheti, Alberto, Paradisi, Federica: Subject indexing in Italy: Recent advances and future perspectives. In: Programme and Proceedings 75th IFLA General Conference and Assembly, Milan, Italy, 23—27 Aug. 2009, <http://www.ifla.org/files/hq/papers/ifla75/200-lucarelli-en.pdf>

тематичен достъп. Примери в това отношение са Китай, Чехия, Естония, Нидерландия, Полша, Италия, Франция, Германия, Финландия, Русия, България и др.

При прегледа на достъпните бази от данни с контролирани речници на националните библиотеки, бе отчетена тенденцията за свързване на различните информационни езици – кодов и словесен. Предметните рубрики се индексират с йерархична класификация. Пример за това можем да посочим в базата данни с нормативни предметни данни на Чешката национална библиотека*, където освен приет термин и логически свързани термини, се посочва и класификационният индекс, на който отговаря предметната рубрика (ил. 7). Обикновено нормативните записи на предметните рубрики са преведени на английски език.

Илюстрация 7. Пример за нормативен запис на предметна рубрика от Нормативната база от данни на Националната библиотека на Чехия.

The screenshot shows the 'Databases of the National Library CR' interface. At the top, there is a navigation bar with 'Logout' and 'Login' links, and a 'Search / Browse' button. Below this, the main content area displays 'AUT - Full View of Record'. It includes a 'Choose format' section with 'Standard' and 'MARC' options, and 'Record 38 out of 146'. A table lists various fields: Control no. (fd993392), Heading (elektronická hudba), Seen from (hudba elektronická), Broader term (elektroakustická hudba), Related term (ambientní hudba), UDC (789.983), and English (electronic music). At the bottom, the System no. is 000993392, and a URL is provided: [https://aleph.nkp.cz/F/?func=direct&doc_number=000993392&local_base=AUT]

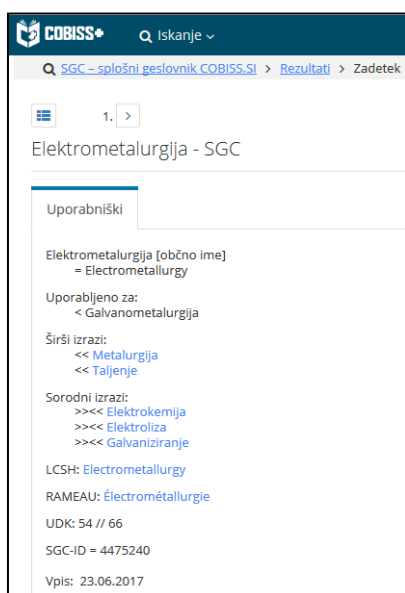
Control no.	fd993392
Heading	elektronická hudba
Seen from	hudba elektronická
Broader term	elektroakustická hudba
Related term	ambientní hudba
UDC	789.983
English	electronic music
System no.	000993392

Друг, добър пример, който илюстрира свързването на различни стандартни инструменти за индексирание е от базата с нормативни данни на словенската национална

* <https://www.en.nkp.cz/>

библиотека^{**}. Всеки нормативен запис, освен типичната за тезаурус структура (приет термин за индексирание, неприет и свързани термини (род-вид и асоциативно), съдържа и еквивалентна рубрика от предметните рубрики на Конгресната библиотека (LCSH) и от „Нормативния предметен списък“ на Френската национална библиотека и библиография (Répertoire d'autorité-matière encyclopédique et alphabétique unifié, RAMEAU), както и индекс по УДК.

Илюстрация 8. Пример за индексирани от словенската национална библиотека.



Приложимост в модела на лингвистично осигуряване на българска електронна библиотека с универсална тематика

Както показва световният опит съществуват национални предметизационни схеми или тезауруси. Те отразяват максимално адекватно и пълно спецификата на местната култура. Националната библиотека „Св. св. Кирил и Методий“ (НБКМ) и българската национална библиография също създават и поддържат общ по тематика тезаурус. Той е създаден на базата на списъка с предметни рубрики на библиотеката, но е развит така, че да отразява родово-видови и асоциативни връзки. Тезаурусът се използва за отразяване на темата на документите, регистрирани от българската национална библиография и за тематичен достъп в електронния каталог на НБКМ. Към настоящия момент тезаурусът се конвертира в MARC формат (COMARC/A за нормативни данни) и се разработва като база данни.

^{**} <https://www.cobiss.net/>

Автоматизирано тематично индексирание

Интелектуалното (ръчно) индексирание се осъществява от индексатори, които използват своите знания и опит, за да определят съдържанието на документите. Подготовката на индексатори е свързано с времеви и финансови ограничения. Същевременно обемът на информацията и документите расте непрекъснато, а това изисква нови работни процедури на тематично индексирание, по-бързи от традиционното индексирание. Компютърните технологии могат да помогнат на индексатора²⁸⁶. Автоматизираното индексирание с информационни езици се осъществява посредством компютърни алгоритми. Обикновено то отчита всички или почти всички думи и позволява многоаспектно представяне съдържанието на документите. В специализираната литература се водят дебати относно предимствата и недостатъците на автоматизираното индексирание. Изследователи отбелязват, че компютърното генериране на индекси все още има ниска степен на точност²⁸⁷. През последните 20 години се осъществяват няколко проекта за автоматизирано индексирание и класифициране²⁸⁸. Въпреки това процесите по автоматичното тематично индексирание търпят развитие и все още не се прилагат като единствен метод.

Напълно автоматично може да се индексира чрез езика на ключовите думи. Като процесът включва математически и статистически методи за изчисляване честотата на срещане на думи в текста, които да бъдат отразени в библиографското описание.

Потребителски интерфейс

Интерфейсът е от съществено значение за реализирането на пълните възможности на информационните езици. Без интерфейсът да е обект на настоящото изследване ще предложим някои основни изисквания за осигуряване на удобна навигация по тема и

²⁸⁶ **Svensson**, Lars G., **Jahns**, Yvonne. PDF, CSV, RSS and other Acronyms: Redefining the Bibliographic Services in the German National Library. In: *Programme and Proceedings* [online]. 76th IFLA General Conference and ssembly, Gothenburg, Sweden, 10–15 Aug. 2010. [Viewed 21.06.2018]. Available from: <https://www.ifla.org/past-wlic/2010/91-svensson-en.pdf>

²⁸⁷ **Shields**, Ginger. *LI-842 Automatic Indexing Assignment*. [Viewed 21.06.2018]. 2005. Available from: http://www.shieldsnetwork.com/LI842_Shields_Automatic_Indexing.pdf

²⁸⁸ *виж например*: Indexing Initiative Project на Националната медицинска библиотека на САЩ, достъпно от: <http://i.nlm.nih.gov/semiauto.shtml>; Проект за автоматизирана многоотраслова класификация проведен от университетите в Лунд и OCLC виж : **Koch**, Vazine-Goetz. *Automatic classification and content navigation support of Web services: DESIRE II cooperates with OCLC*; **Yi**, Kwan. *Challenges in automatic classification using library classification schemes*. [Viewed 21.06.2018]. Available from: <http://archive.ifla.org/IV/ifla72/papers/097-Yi-en.pdf>; **Moeller**, G. et al. Automatic classification of the World-Wide-Web using UDC. In: *Proceedings of the GfKI*. Heidelberg : Springer, 1999 **Larson**, Ray L. Experiments in automatic Library of Congress Classification In: *Journal of the American Society for Information Science*. 1992, № 43, p. 130–148; CADIAL project. [Viewed 21.06.2018]. Available from: <http://www.cadial.org/index.php?>

ефективно търсене по елементите на библиографското описание данни в електронна библиотека:

1) Интерфейсът трябва да използва пълния потенциал на библиографското описание и да е възможно търсене по всички важни за идентифицирането на документите елементи.

2) Препоръчително е интерфейсът да осигурява възможност както за основно търсене в едно поле, така и разширено търсене с възможност за комбиниране на признаците за търсене чрез логически оператори.

3) Потребителите трябва да имат възможност да конкретизират (филтрират) резултатите от търсенето по различни признаци (фасети), освен по автор, език, жанр, година, е полезно да има възможност за ограничаване на резултатите от търсенето по тема, географска рубрика, хронологична рубрика или да се прецизира по дисциплина посредством списък на систематични индекси.

4) Предметните индекси трябва да бъдат включени във формата за показване на библиографското описание. Това дава на потребителя вход към съдържанието на документа и е отправна точка за следващо търсене по тема.

5) Електронната библиотека трябва да осигурява показване и преглеждане на информационния масив по тема като използва високите нива на йерархичните класификации или широки тематични категории. Така създадената логическа структура на знанието позволява на потребителите да намерят релевантни документи леко и бързо, дори без да са ориентирани в предметната област на електронната библиотека.

6) Осигуряване на достъп на потребителите до нормативни данни. Възможността за избор на индекс от контролиран речник позволява на потребителите да подберат точно търсената тема (понятие). Много полезно е да се показват не само приетите термини, но и еквивалентните и свързаните термини. При класификационни схеми трябва да може да се преглежда йерархичната структура, да се виждат кодовите означения и да е предвидена възможност за търсене по словесния израз на индексите.

7) Интерфейсът трябва да позволява многоезичен достъп като тук имаме предвид главно възможността за свързване на речници на различни информационни езици. Възможно е използването на програма, която да превежда термина от един информационен език на друг, за да подпомогне търсенето по тема.

С оглед разнообразието от потребители, електронната библиотека трябва да предлага различни методи и средства за търсене. Наред с информационните езици трябва да се осигури възможност за пълнотекстово търсене. То се основава на различни

математически и статистически методи и поради това, че се осъществява на естествен език не е обект на настоящото изследване.

Изводи

1. Обемът и съдържанието на понятието „лингвистично осигуряване” зависи от обекта и сферата на неговото приложение. Лингвистичната осигуреност на електронна библиотека се базира на използването на комплекс от информационни езици – класификационни и словесни, с различно предназначение и роля. Те алгоритмизират и автоматизират обработката на метаинформацията на документите, което позволява постигането на прецизно индексирание и на по-пълно и релевантно търсене.

2. Основателният избор на информационен език трябва да отчита и неговата структура – азбука, лексика и отношения между лексикалните единици, граматика. Необходимо е структурата да позволява съставяне с голяма точност както на търсещия образ на документа, така и на запитването. Задължително е информационният език да притежава съответната лексикално–семантична база, която да е в състояние да опише спецификата на информационните обекти в конкретната електронна библиотека. За постигане качество на индексирането и на търсенето, минимизиране на неизбежната загуба на информация при него, както и за улеснение на потребителя, правещ запитване към системата, съществено важен е оптималният състав на речника от понятия, тяхната лексикографска и семантична обработка.

3. Лингвистичното осигуряване на изследваните български електронни библиотеки не включва комплекси от информационни езици, които едновременно да предоставят възможност за цялостно представяне характеристиките, предмета (темата) на документите и многоаспектното разкриване на тяхното съдържание, както и за осъществяване на различни видове търсене. Наложително е приобщаването на наложените в практиката на традиционните библиотеки езикови средства с използването на контролирани речници - йерархична класификационна схема по области на знанието, предметна класификация, дескрипторен език.

4. Автоматизираното индексирание с информационни езици, което се осъществява посредством компютърни алгоритми все още не е достатъчно развито, така че да притежава удовлетворяваща степен на точност. Напълно автоматично може да се индексира чрез езика на ключовите думи.

5. В световен мащаб не съществува общ стандарт за тематично индексирание, е препоръчително използването на широко разпространените международни класификации за тематичен достъп и търсене в библиотеките. По този начин се осигурява възможност за международен обмен на данни, интегриране на различни метаданни и колекции и потенциален многоезичен достъп. Както УДК, така и съкращеният и вариант UDC Summary може пълноценно да се използват за индексирание на съдържанието на документите и информационните запитвания и за структуриране и систематизиране на информационните обекти в електронна библиотека.

6. В лингвистичната осигуреност на електронната библиотека е препоръчително включването на естествен език, съответстващ на международни принципи и стандарти, като ефективен инструмент за словесен тематичен достъп. Обоснован избор за България е създаденият и поддържан от Националната библиотека „Св. св. Кирил и Методий” и Българската национална библиография общ по тематика тезаурус. Той отразява адекватно и максимално пълно спецификата на нашата култура.

ЗАКЛЮЧЕНИЕ

Електронната библиотека притежава характеристиките на традиционната библиотека, като добавя предимства, които я превръщат в надеждно средство за съхраняване и ефективно използване на разнообразни колекции от документи (текст, видео, аудио и др.) с унифициран подход към тяхната организация и достъп. Нейната архитектура и програмна среда, както и методите за подбор и съхранение гарантират качеството на информацията и услугите и осигуряват достъп и ефективни средства и методи за търсене и навигиране на информация.

За да се постигне широко и многократно използване на съдържание, както и взаимодействие с други информационни системи е необходимо спазването на утвърдени спецификации, стандарти и модели за представяне на информацията, особено постигането на семантична оперативна съвместимост, позволяваща формулирането и изпълнението на сложни информационни запитвания, подобряване точността на търсене, интегриране на хетерогенни информационни ресурси.

За да се преодолеят затрудненията в търсенето и обмена на информация се развиват все повече инструменти и решения, които целят да доведат до възможността машините да „разбират“ естествения език. Създават се модели за формализирано представяне и използване на знания и на методи за тяхната машинна обработка. лексикални семантични ресурси. Те могат да бъдат използвани за концептуално индексирание и търсене в електронните библиотеки. Организацията на знанието е тясно свързана с библиотечно-информационните науки, като в тесния си смисъл има най-изявена връзка с класификационните на информационните ресурси, тяхното описание и индексирание. Системите за организация на знания (като напр. семантични мрежи, онтологии, метаданни, тезауруси и мн.др.) улесняват управлението и търсенето на информация и знания от хора и машини. Семантичният уеб, или уеб 3.0 се разработва с цел да постигне именно разбирането на естествения език от машините, обработващи информацията. По този начин ще се осигури на потребителя възможност за семантично търсене, улесняване на достъпа до големи обеми от данни, многократното използване и интерпретиране в зависимост от целите. Изграждането е въз основа на уеб стандарти, метаезици, онтологии и логически правила. Развитието на електронната библиотека, интерпретирана като информационна системи, се свежда все повече до повишаване на стандартизацията на отделните процеси и процедури, както и до въвеждането на норми, относно обработката на обектите в електронната библиотека. Колкото повече и по-

качествена и релевантна информация целим да организираме в една електронна библиотека, толкова по-стандартизирани и предварително контролирани (нормализирани) входове в системата трябва да осигурим. За да постигнем тази зависимост е необходимо задълбочено и последователно планиране и изграждане на всички компоненти на електронната библиотека, и особено нейното лингвистично осигуряване.

Лингвистичната осигуреност на електронна библиотека се базира на използването на комплекс от информационни езици – класификационни и словесни, с различно предназначение и роля и позволяващи прецизно индексирание. При изборът на информационен език трябва да се отчита неговата структура, лексикално-семантична база, оптималния състав на речника. Подборът на информационните езици трябва да се основава на техните най-съществени характеристики.

Въз основа направените наблюдения, анализи и проведеното емпирично изследване на български електронни библиотеки се установи, че трудно може да се говори за лингвистична осигуреност, която да позволи пълноценното им функциониране. Липсват комплекси от информационни езици, които да предоставят възможност за цялостно представяне характеристиките, предмета на документите и многоаспектното разкриване на тяхното съдържанието и осъществяване на различни видове търсене. Наложително е приобщаването на наложените в практиката на традиционните библиотеки езикови средства с използването на контролирани речници.

Наложената практика показва, че автоматизираното индексирание с информационни езици, осъществявано посредством компютърни алгоритми все още не е достатъчно развито, така че да притежава удовлетворяваща степен на точност. За електронна библиотека с универсална тематика е потребна обща класификация, която да обхваща множество понятия от всички отрасли на знанието и човешката дейност. Тъй като в световен мащаб не съществува общ стандарт за тематично индексирание, е препоръчително на национално ниво използването на широко разпространените международни класификации за тематичен достъп и търсене в библиотеките. По този начин се осигурява възможност за международен обмен на данни, интегриране на различни метаданни и колекции и потенциален многоезичен достъп. УДК е подходящ инструмент за индексирание на съдържанието на документите и информационните запитвания, както и за структуриране и систематизиране на информационните обекти в електронна библиотека.

Включването на естествен език в лингвистичното осигуряване на електронна библиотека е ефективен инструмент за словесен тематичен достъп. Обоснован избор за България е създаденият и поддържан от Националната библиотека „Св. св. Кирил и Методий“ и Българската национална библиография общ по тематика тезаурус, който отразява адекватно и максимално пълно спецификата на нашата култура.

В резултат на изпълнението на поставените в изследването задачи се дефинираха същността и функциите на електронната библиотека, като се представиха и съвременните концепции за електронна библиотека. Разглеждани бяха концептуалния модел на електронна библиотека и информационни модели за съдържанието. Изследва се и се описва същността на информационните езици. Представи се ролята и предназначението на информационните езици в състава на лингвистичното осигуряване на електронната библиотека. Проведено би емпирично изследване на използваните информационни езици в българските електронни библиотеки. Разработи се модел на лингвистично осигуряване на българска електронна библиотека с универсална тематика.

В хода на изследването бе потвърдена работната хипотеза, че моделът на лингвистично осигуряване на електронната библиотека трябва да съдържа необходим и достатъчен комплекс от информационни езици, за да изпълнява основните си задачи и функции. Те трябва да гарантират максимално резултативно търсене, идентификация на информационните обекти и взаимодействие с други електронни библиотеки и автоматизирани системи.

Настоящата разработка въвежда идеи за използване на конкретни информационни езици в технологията на електронната библиотека. Те осигуряват различни методи за търсене на документи – по характеристиките на документите, по техния предмет/и (теми), по думи от текста на документа или свързани семантично с тях. На основа на постигнатите резултати бъдещите изследвания могат да бъдат насочени към надграждането на тези методи и средства със семантични технологии и уеб услуги, които ще превърнат електронната библиотека в семантична електронна библиотека. Семантичната електронна библиотека използва онтологии за извличане, съхраняване и използване на знанията. Това предполага подобряване на качеството на търсене в електронната библиотека.

СПИСЪК НА ИЗПОЛЗВАНАТА ЛИТЕРАТУРА

На кирилица:

1. **Антопольский**, А. Б. *Лингвистическое обеспечение электронных библиотек*. Москва: Информрегистр, 2002.
2. **Атанасова**, В. *Ръководство за предметизация на литературата в библиотеките*. София: НБКМ, 1972, 88 с.
3. **Атанасова**, Ирена. *Създаване на експертни системи* [онлайн]. Благоевград: Унив. изд. „Неофит Рилски“, 2018. [Viewed 20.06.2018]. PDF, 4MB. Available from:
https://www.researchgate.net/profile/Irena_Atanasova2/publication/323341660_Szdavane_na_ekspertni_sistemi_Expert_systems_delelopment/links/5a8eb5e50f7e9ba42966fe0d/Szdavane-na-ekspertni-sistemi-Expert-systems-delelopment.pdf
4. **БДС ISO 23081-2:2015** *Информация и документация*. Управление на метаданни за документи/записи. Част 2: Концептуални въпроси и прилагане. София: БИС.
5. **БДС ISO 25577:2015**. *Информация и документация*. MarcXchange. В: *Български институт по стандартизация* [онлайн]. [Прегледан 20.06.2018]. Достъпно от: http://www.bds-bg.org/bg/standard/?natstandard_document_id=74166
6. **БДС ISO 5127** *Информация и документация*. Основни понятия и речник. София: БИС, 2009.
7. **Берестова**, Т. Ф. *Електронна библиотека как инструмент за създаване на единично информационно пространство*. В: *Вестник Челябинской государственной академии культуры и искусств* [онлайн], 2009, 4. [Прегледан 24.05.2018]. Достъпно от: <https://cyberleninka.ru/article/n/elektronnaya-biblioteka-kak-instrument-sozdaniya-edinogo-informatsionnogo-prostranstva> .
8. **Българска академия на науките**. *Официален правописен речник на българския език*. София: Просвета, 2012, 676 с.
9. **Габровска**, Свободозаря, **Пейчева**, Веселина. *Методични и организационни проблеми при съставянето на тезауруси в обществените науки*. В: *Съвременни проблеми на научната информация*. София: БАН, 1974., с. 29–60.
10. **Гай**, Ф. *Развитие на електронни библиотеки: опит на някои национални библиотеки в Северна Америка, Австралия, Азия и Европа*. В: *Електронни библиотеки*, 2000, Т. 3, вып. 3, с. 7–10.
11. **Гендина**, Н. И. *Лингвистически средства на библиотечно-информационни технологии*. Санкт-Петербург: Професия, 2015.
12. **Георгиев**, Веселин. *Насоки за развитие на корпоративните комуникации чрез приложение на средства от социалните мрежи*. В: *Икономически и социални алтернативи* [online]. 2014, 3, с. 127–137. [Прегледан 16.06.2018]. Достъпно от: <https://www.unwe.bg/alternativi/bg/journalissues/article/5881>
13. **Гергова**, Ани, състав. *Енциклопедия Българска книга*. София: Пенсофт, 2004.
14. **Глинков**, Сергей С., **Клисурова**, Марина, **Радева**, Румяна. *Обзор на ресурсите с отворен достъп в България* [online]. [Прегледан 21.04.2018]. Достъпно от:

- https://www.researchgate.net/publication/320827255_A_Review_of_the_Current_Situation_of_Open_Access_in_Bulgaria
15. **Голицына, О. Л., Максимов, Н. В.** Технология и средства доступа к распределенным ресурсам электронных библиотек и баз данных. В: *Электронные библиотеки*, 2000, Т. 3, вып. 1, с. 16–23.
 16. **Горный, Евгений, Вигурский, Константин.** *Развитие электронных библиотек: Мировой и российский опыт, проблемы, перспективы* [онлайн]. PDF, 294.19 KB. [Прегледан 12.12.2016]. Достъпно от: <http://www.ifap.ru/library/book004.pdf>
 17. **Делийска, Боряна.** *Електронно обучение, базирано на онтологии, в областта на компютърните мрежи и екоинформатиката: Хабилитационен труд* [online]. София, 2008. [Прегледан 20.06.2018]. Достъпно от: https://www.researchgate.net/publication/319507476_Elektronno_obucenie_bazirano_na_ontologii_v_oblastta_na_komputrните_mrezi_i_ekoinformatikata
 18. **Дигитални библиотеки с мултимедийно съдържание и приложения в българското културно наследство** [онлайн]. Изследователска разработка по договор ИД 8/21.07.2005 с Агенция „Развитие на съобщенията и на информационните и комуникационните технологии”. I етап. [Прегледан на 13.12.2016]. Достъпно от: http://mdl.cc.bas.bg/Digital_libraries_with_multimedia_content_and_applications_in_Bulgarian_cultural_heritage.pdf
 19. **Дипчикова, Александра, Тотоманова, Антоанета.** *Промените в системите за организация на знанието – предизвикателство към българската библиотечна общност* [презентация]. XXI национална годишна конференция на ББИА, 10 юни 2011, Благоевград. [Прегледан 15.06.2018 г.]. Достъпно от: www.lib.bg/konferencii1/nk2011/prezentacii/atotomanova_adipchikova.ppt
 20. **Дипчикова, Александра, Тотоманова, Антоанета.** УДК – инструмент за оперативна съвместимост на достъпа по съдържание в многоезична мрежова среда [online]. В: *Годишник на Българския информационен Консорциум* [онлайн]. 2013, с. 17–21. [Прегледан 15.06.2018 г.]. Достъпно от: <https://www.bic.bg/files/booklet.pdf>;
 21. **Дипчикова, Александра, Тотоманова, Антоанета.** Универсална десетична класификация – съвременни направления в развитието и прилагането им в България. В: *Библиотека*, 2013, № 2–3, с. 99–108.;
 22. **Добрева, Стефка.** Социално търсене в библиотечни бази от данни чрез използване на системата REDIS. В: *Научни трудове на Русенския университет* [online]. 2014, том 53, серия 3.2, с. 138–143 [Прегледан 20.05.2018]. Достъпно от: <http://conf.uni-ruse.bg/bg/docs/cp14/3.2/3.2-24.pdf>
 23. **Дончева, Анета.** *Видове библиотеки при промяна на парадигмата на библиотечната теория и практика*. София: Нац. библ. „Св. св. Кирил и Методий“, 2015.
 24. **Ершова, Т. В., Хохлов, Ю. Е.** Межведомствена програма „Российские электронные библиотеки“: Подходи и перспективи. В: *Электронные библиотеки*, 1999, Т. 2, Вып. 2, с. 29–33.
 25. **Желев, Янислав, Тошков, Ангел, Николова, Евгения, Желева, Мария.** Развитие на икономика, базирана на знания и иновации В: *Бизнес посоки*

- [онлайн]. 2015, 1, с. 88–99. [Прегледан 06.07.2017]. Достъпно от: http://research.bfu.bg:8080/jspui/bitstream/123456789/617/1/BFU_BP_2015_1_Zheleva_Toshokov_Nikolova_Zheleva.pdf
26. **Закон** за висшето образование. В: *Lex.bg* [online]. [Прегледан 25.05.2018]. Достъпно от: <https://lex.bg/bg/laws/ldoc/2133647361>.
 27. **Земсков, А. И., Шрайберг, Я. Л.** *Электронная информация и электронные ресурсы*. Москва: Фаир, 2007.
 28. **Иванова, Татьяна Ивановна.** *Анализ и разработка на семантични технологии за търсене в уеб: Дисертация*. София: ТУ-София, 2009.
 29. **Иванова, Татьяна.** *Глобалната семантична мрежа: Технологии, инструменти, приложения*. София: Технически унив., 2015.
 30. **ИФЛА.** ISBD Международният стандарт за библиографско описание. Консолидирано издание. София: Българска библиотечно-информационна асоциация, 2012. Достъпен и от: <https://www.ifla.org/files/assets/hq/publications/series/44-bg.pdf>
 31. **ИФЛА.** Изложение на международните принципи на каталогизацията. 2009.
 32. **Камбуров, Кирил Аспарухов.** *Сравнителен анализ на онлайн платформи за електронни периодични издания: Дисертация*. София, 2018, с. 78.
 33. **Когаловский, М. Р.** Метаданни в компютърних системах. В: *Програмиране* [онлайн], 2013, № 4, с. 28-46. [Прегледан на 24.05.2018]. Достъпно от: <http://www.ipr-ras.ru/articles/kogalov13-03.pdf>
 34. **Когаловский, М. Р.** *Перспективныe технологии информационных систем*. Москва : ДМК, 2003.
 35. **Когаловский, М. Р.** Развитие стандартов XML: Новые возможности и применения [online]. [Прегледан на 24.05.2018]. Достъпно от: <http://www.ipr-ras.ru/articles/conf02st.htm>
 36. **Коева, Светла.** WordNet и BulNet. В: *Езикови ресурси и технологии за български език*. София: Акад. изд. „Проф. Марин Дринов“, 2014, с. 154–173.
 37. **Коева, Светла.** *Предговор*. В: *Езикови ресурси и технологии за български език*. София: Акад. изд. „Проф. Марин Дринов“, 2014.
 38. **Козлов, Д. Д.** *Информационно-поисковые системы в Internet* [онлайн]: Текущее состояние и пути развития: Технологический обзор. Москва: Московский Государственный Университет им. М. В. Ломоносова, 2000. [Прегледан 06.07.2017]. PDF, 408 KB. Достъпно от: http://lvk.cs.msu.su/~ddk/ir_and_ia_review.pdf
 39. **Койчева, Елена.** The Dublin Core Metadata Initiative стандартизиран формат за метаданни за интернет пространството. В: *Computer*, 2004, № 12, с. 55–58.
 40. **Койчева, Елена.** Информационна евристика: систематичен план за провеждане на информационно търсене в уеб ресурси. В: *Годишник на Софийския университет „Св. Климент Охридски“*. Философски факултет. Библиотечно-информационни науки, 8, 2016, с. 287–303. Достъпно онлайн на: <https://research.uni-sofia.bg/handle/123456789/1404>

41. **Костова**, Десислава, **Минкова**, Соня. *Библиотекознание и теория на научната информация*. София: ВМГИ, 1990.
42. **Лойко**, В. И., ред. *Информационные системы и технологии в экономике* [онлайн]. Москва: Финансы и статистика, 2005. [Прегледан 12.03.2018]. PDF, 19.1 МВ. Достъпно от: <http://lib.maupfib.kg/wp-content/uploads/2015/12/end/akademy/03%20inf%20sis%20i%20teh%20v%20econ/inf%20sis%20i%20teh%20v%20econ%20loiko.pdf>
43. *Манифест на ИФЛА/ЮНЕСКО за дигиталните библиотеки* [онлайн]. [Прегледан 12.12.2016]. Достъпно от: <http://www.ifla.org/files/assets/digital-libraries/documents/ifla-unesco-digital-libraries-manifesto-bg.pdf>
44. **Миланова**, Милена. Библиографски данни в семантичния веб. В: *Годишник на Софийския университет „Св. Климент Охридски“*. Философски факултет. *Книга библиотечно-информационни науки* [online]. Т. 7, 2015, с. 201–214. https://research.uni-sofia.bg/xmlui/bitstream/handle/123456789/1381/7_Milena%20Milanova.pdf?sequence=1
45. **Миланова**, Милена. *Българската каталогизация в глобалното информационно пространство на XXI век. Анализи, стратегии, перспективи: Дисертация* [online]. София, 2008 [прегледан на 27.06.2018]. Достъпно от: https://research.uni-sofia.bg/bitstream/10506/1089/5/Milana%20Milanova_diss.pdf
46. **Михайлов**, А. И., **Черный**, А. И., **Гиляревский**, Р. С. *Основы информатики*. Москва: Наука, 1968, с. 757.
47. **Младенова**, Мария. *Библиотечни класификации* [online]. [София: Полувисш инст. по библ. дело], 1996. [Прегледан 20.06.2018 г.]. Достъпно от: https://bibliographylibrary.weebly.com/uploads/3/7/9/4/37944541/bibliotechni_klasifikacii.pdf
48. **Наков**, Светлин. Автоматично извличане на фалшиви приятели от паралелен двуезичен корпус [онлайн]: Автореферат на дисертационен труд. София, 2009. [Прегледан 28.02.2018]. Достъпно от: <https://www.nakov.com/wp-content/uploads/2010/05/nakov-phd-avtoreferat-false-friends.pdf>
49. **Николов**, Румен. *Глобалният кампус* [online]. София: Авангард принт, 2009 [Прегледан 26.05.2018]. Достъпно от: www.bvu-bg.eu/R.Nikolov-Globalniat_kampus.pdf
50. **Нишева**, Мария М., **Шишков**, Димитър П. *Изкуствен интелект*. Добрич: Интеграл, 1995.
51. **Нишева-Павлова**, Мария. *Създаване на хетерогенни цифрови библиотеки, основано на онтологии: Хабилитационен труд* [онлайн]. София, 2013. [Прегледан 22.05.2017]. Достъпно от: https://www.fmi.uni-sofia.bg/sites/default/files/habilitation_work_of_professor/habilit_trud_mn.pdf
52. **Ноев**, Николай Генчев. *Подходи и методи за създаване, съхранение, представяне и защита на цифрови ресурси от областта на културно-историческото наследство чрез технологии, основани на знания: Дисертация* [online]. София: БАН, 2015. [Прегледан 20.06.2018]. Достъпно от: <http://www.math.bas.bg/vt/nnoev/phd/files/phd-thesis-NN-2015-bg.pdf>

53. **НЭБ**: Национална електронна библиотека [онлайн]. [Прегледан 28.05.2018 г.]. Достъпно от: <https://нэб.рф/about/>
54. **Оророва**, Даниела. Извличане на знания от данни в образователно пространство. В: *Годишник на БСУ* [online]. XXXIV, 2016, с. 5–19 [Прегледан 25.05.2018]. Достъпно от: <http://research.bfu.bg:8080/jspui/handle/123456789/650>
55. **Павлов**, Радослав, **Панева-Маринова**, Десислава. Цифровизация и цифрови библиотеки – образователни аспекти. В: *Образованието в информационното общество: Сборник доклади на Национална конференция* [онлайн]. 26–27 май 2011, Пловдив. София: Асоциация „Развитие на информационното общество”, 2011. [Прегледан 09.07.2017]. ISSN 1314-0752. Достъпно от: <https://core.ac.uk/download/pdf/62660570.pdf>
56. **Павлов**, Радослав. Цифрови библиотеки с културно-историческо и научно съдържание [онлайн]. [Прегледан 14.05.2017]. Достъпно от: mdl.cc.bas.bg/radko/rp/Pavlov-lecture_final.ppt
57. **Павлов**, Радослав. *Цифрови библиотеки с научно съдържание – реализации и проекти* [онлайн]. [Прегледан на 14.05.2017]. Достъпен от: raspl.org/wp-content/uploads/2014/02/10_R.Pavlov-DL-Science.pptx
58. **Павлова**, Лилия Радославова. *Семантични технологии и средства в цифрови библиотеки за художествено културно наследство: Автореферат* [online]. София, 2013, с. 13 [Прегледан 20.06.2018]. Достъпно от: http://www.math.bas.bg/IMIdocs/ZRASRB/docs/Lilia_Pavlova/LPavlova-avtoreferat.pdf
59. **Павловска**, Елена. От библиотеката на Ярослав Мъдри до националната електронна библиотека на Русия. В: *През призмата на информационната наука*. София: Унив. изд. „Св. Климент Охридски”, 2016, с. 329–339.
60. **Попова**, М. *Теория на терминологията*. Велико Търново, 2012.
61. **Рашидов**, Алдениз, **Ненов**, Тошко. *Оценяване на интелигентността на информационни системи* [онлайн]. January 2006. [Прегледан 24.06.2018]. Достъпно от: https://www.researchgate.net/publication/270161120_OCENAVANE_INTELLIGENT_NOSTTA_NA_INFORMATIONNI_SISTEMI_Estimation_of_the_degree_of_Intelligence_of_information_systems
62. *Ролята на Европа за цифровия достъп, видимостта и използването на европейското културно наследство – Заключение на Съвета (31 май 2016 г.)* [онлайн]. [Прегледан 07.07.2017]. Достъпно от: <http://data.consilium.europa.eu/doc/document/ST-9643-2016-INIT/bg/pdf>
63. **Руководство** по обеспечению тематического доступа в национальной библиографии. В: *IFLA* [онлайн]. Санкт-Петербург, 2017. [Прегледан 20.06.2018] Достъпно от: <https://www.ifla.org/files/assets/hq/publications/series/45-ru.pdf>
64. **Семантичната** мрежа на балканските езици е част от световната WordNet. В: *Computerworld* [онлайн]. 2 март 2006. [Прегледан 28.02.2018]. Достъпно от: http://computerworld.bg/11835_semantichnata_mrezha_na_balkanskite_ezici_e_chast_ot_svetovnata_wordnet
65. **Стефанов**, Красен, **Бойчев**, Павел, **Стефанова**, Елиза, **Георгиев**, Атанас, **Николова**, Николина, **Григоров**, Александър. Цифровите библиотеки в

- обучението на учители [online]. В: *Mathematics and education in mathematics: Proceedings of the Fortieth Jubilee Spring Conference of the Union of Bulgarian Mathematicians Borovetz, April 5–9 2011*. PDF, 168.25 KB. Достъпно от: www.math.bas.bg/smb/2011_PK/tom/pdf/120-135.pdf
66. **Стефанова**, Радослава Марио. *Електронните ресурси с отдалечен достъп и националната библиография: Дисертация*. София, 2015.
 67. **Столяров**, Ю. Н. Библиотека – двухконтурная система. В: *Научные и технические библиотеки*, 2002, № 11, с. 9.
 68. **Сюнтюренко**, О. В., **Хохлов**, Ю. Е. Распределенные библиотечные сети и электронные библиотеки в России. В: *Электронные библиотеки*, 2000, Т. 3, Вып. 5, с. 6–10.
 69. **Терзиева**, Годорка, **Тотков**, Георги. Един онтологичен подход към процеса на обучение. В: Национална конференция Образованието в информационното общество [online]. 12–13 май 2009, Пловдив. София: Асоц. Развитие на информ. о-во, 2009, с. 17–24. [Прегледан 20.06.]. Достъпно от: <http://sci-gems.math.bas.bg:8080/jspui/handle/10525/1358>
 70. **Тодорова**, Радостина. *Свободен достъп до научни електронни публикации в Нов български университет* [online]. [Прегледан 23.05.2018]. Достъпно от: https://www.researchgate.net/publication/271450395_Svoboden_dostp_do_naucni_elektronni_publicacii_v_Nov_blgarski_universitet_6_Open_access_to_scientific_publications_in_the_New_Bulgarian_University
 71. **Тотоманова**, Антоанета. *Стандартизация на класификационната система в България* [презентация]. Национална кръгла маса „Стандартизацията в библиотеките“ 20–21 окт. 2016, София. [Прегледан 15.06.2018 г.]. Достъпно от: https://www.lib.bg/kragli_masi/atotomanova-2016.ppt;
 72. **Тотоманова**, Антоанета. Структура и достъп до българската текуща национална библиография в контекста на системата COBISS. В: *Globalization and the Management of Information Resources* [онлайн]. Papers from the International Conference Sofia, Bulgaria, 12–14 November 2008, с. 110–117. Достъпно от: https://research.uni-sofia.bg/xmlui/bitstream/handle/10506/455/SOFIA2008_new.pdf?sequence=3
 73. **Тузовский**, А. Ф., **Чириков**, С. В., **Ямпольский**, В. З. *Системы управления знаниями (методы и технологии)*. Томск: Изд-во НТЛ, 2005.
 74. **Търнина**, Албена. Технологиите на семантичния уеб в услуга на електронното обучение. В: *Списание на Софийския университет за електронно обучение* [online]. 2011, 1. [Прегледан 20.06.2018]. Достъпно от: <https://journal.e-center.uni-sofia.bg/f/downloads/2011/Broi%201/A.Turnina.pdf>
 75. *Универсална десетична класификация: Съкратено издание*. София: Нац. библ. „Св. св. Кирил и Методий“, 2017, 446 с.
 76. **Федотов**, А. М., **Баракнин**, В. Б., **Жижимов**, О. Л., **Федотова**, О. А. Модель информационной системы для поддержки научно-педагогической деятельности. В: *Вестник НГУ. Серия Информационные технологии*, 2014, Т. 12, вып. 1.
 77. *Функционални изисквания към библиографските записи: Концептуален модел* [online]. [Прегледан 06.04.2017]. PDF. Достъпно от: https://www.ifla.org/files/assets/cataloguing/frbr/frbr_2009-bg.pdf; **Functional**

- requirements for bibliographic records: Final report* [online]. München: K.G. Saur, 1998. As amended and corrected through February 2009. [Viewed 17.07.2017]. PDF. Available from: http://www.ifla.org/files/assets/cataloguing/frbr/frbr_2008.pdf
78. **Функционални** изисквания към нормативните данни: *Концептуален модел* [online]. [Прегледан 24.05.2018]. PDF. Достъпно от: https://www.ifla.org/files/assets/cataloguing/frad/frad_2009-bg.pdf; **Functional requirements for authority data: A conceptual model** [online]. As amended and corrected through July 2013. [Viewed 24.05.2018]. PDF. Available from: http://www.ifla.org/files/assets/cataloguing/frad/frad_2013.pdf
79. **Харизанова, Оля.** *Библиотеките и мрежовото общество: Ефекти и трансформации.* София: Унив.изд. „Св. Климент Охридски“, 2010.
80. **Харизанова, Оля.** Новите измерения на World Wide Web и предизвикателствата пред библиотеките [online]. В: *Годишник на Софийския университет „Св. Климент Охридски“.* Философски факултет. Книга библиотечно-информационни науки [онлайн]. 2, 2010, с. 123–162. [Прегледан 20.05.2018]. Достъпно от: https://research.uni-sofia.bg/bitstream/10506/316/1/ONharizanova_godishnik.pdf
81. **Хоай, Ле.** *Исследование и разработка электронных библиотек на основе явного описания семантики ресурсов с использованием технологий Semantic Web: Диссертация.* Томск, 2013, с. 34.
82. **Христова, Боряна, Панчева, Цветанка, Чакалов, Андрей.** *Опитът на Националната библиотека „Св. Св. Кирил и Методий“ в дигитализирането на книжния фонд* [online]. [Прегледан 09.07.2017]. Достъпно от: <http://www.bvu-bg.eu/sem/sem8/VHristova.pdf>.
83. **Цокева, Весела.** Семантичната мрежа – от намиране на информация към получаване на знания. В: *През призмата на информационната наука.* София: Унив. изд. „Св. Климент Охридски“, 2016, с. 175–192.
84. **Чавдаров, Станко.** *Теория и методи на научната информация: Обработване, търсене, разпространяване.* София: Наука и изкуство, 1984, 151 с.
85. **Шарон, Т., Франк, А.** Електронни библиотеки в Интернет: Доклад 029-142 на 66-й Генерална конференция ИФЛА, 13–18 авг. 2000 г., Иерусалим, Израил. В: *Научные и технические библиотеки*, 2001, № 2.
86. **Шокин, Ю. И., Федотов, А. М., Барахнин, В. Б.** *Проблемы поиска информации.* Новосибирск: Наука, 2010.
87. **Яврукова, Биляна.** *Библиографски контрол в процеса на ретроспективна конверсия и дигитализация на фондовете: Дисертация.* София, 2015.
88. **Янакиева, Елена.** Нови системи за обединен достъп до информация. В: **Лозанова-Белчева, Елица,** състав. *През призмата на информационната наука.* София: Унив. изд. „Св. Климент Охридски“, с. 193–209.

На латиница:

89. **043 Geographic Area Code (NR).** In: *OCLC* [online]. Support & Training. [Viewed 21.06.2018]. Available from: <https://www.oclc.org/bibformats/en/0xx/043.html>

90. **651 - Subject Added Entry-Geographic Name (R)**. In: *MARC 21 Bibliographic - full* [online]. December 2017. [Viewed 21.06.2018]. Available from: <http://www.loc.gov/marc/bibliographic/bd651.html>
91. A **virtual** library for 8 UEMOA countries under discussion in Dakar. In: *UNESCO* [online]. UNESCO Office in Dakar, 14 April 2014. [Viewed 24.05.2018]. Available from: http://www.unesco.org/new/en/dakar/about-this-office/single-view/news/a_virtual_library_for_8_uemoa_countries_under_discussion_in/
92. **About** DSpace@MIT [online]. In: *MIT Libraries* [online]. [Viewed 25.05.2018]. Available from: <https://libguides.mit.edu/dspace>
93. **About** the Library. In: *Digital.Bodleian* [online]. [Viewed 25.05.2018]. Available from: <https://digital.bodleian.ox.ac.uk/about.html>
94. **About**: EuropeanaConnect is a Best Practice Network funded by the European Commission within the area of Digital Libraries of the eContentplus Programme. In: *EuropeanaConnect* [online]. [Viewed 24.05.2018]. Available from: <http://www.europeanaconnect.eu/about.php>
95. **Anderson**, James. D., **Pérez-Carballo**, José. The nature of indexing: How humans and machines analyze messages and texts for retrieval. In: *Information Processing & Management*, 2001, 37, p. 231–277.
96. **ANSI/NISO Z39.19-2005 (R2010)** Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies;
97. **Appendix B** – Full Level Record Examples. In: *MARC 21 Classification* [online]. January 2000. [Viewed 21.06.2018]. Available from: <https://www.loc.gov/marc/classification/examples.html#lcche>
98. **Arms**, William Y. *The Digital Libraries*. Cambridge, Massachusetts: MIT Press, 2000.
99. **Artale**, A., **Franconi**, E., Guarino, N., **Pazzi**, L. Part-Whole Relations in Object-Centered Systems: An Overview. In: *Data and Knowledge Engineering* [online]. Vol. 20, 1996, 3, p. 347–383 [Viewed 20.06.2018]. Available from: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.33.131&rep=rep1&type=pdf>
100. **Bainbridge**, David, **Nichols**, David M. *Digital Libraries in Education: Specialized Training Course: Study Guide* [online]. Moskow: Unesco Institute For Information Technologies In Education, 2006. [Прегледан 06.07.2017]. PDF, 1.1 MB. Достъпно от: <http://iite.unesco.org/pics/publications/en/files/3214563.pdf>
101. **BalkaNet** – Design and Development of a Multilingual Balkan WordNet. In: *Balkanet Word* [online]. [Viewed 20.06.2018]. Available from: <http://www.dblab.upatras.gr/balkanet/index.htm>
102. **Bangemann** report: Europe and the global information society. In: *European Commission* [online]. CORDIS. Last updated on: 1994-07-07. Record Number: 2730. [Viewed 16.04.2017]. Available from: http://cordis.europa.eu/news/rcn/2730_en.html
103. **Bekiari**, Chryssoula, **Doerr**, Martin, **Bœuf**, Patrick Le, **Riva**, Pat, eds. *Definition of FRBRoo: A conceptual model for bibliographic information in object-oriented formalism* [online]. Version 2.4. November 2015. [Viewed 24.05.2018]. PDF. Available from: https://www.ifla.org/files/assets/cataloguing/FRBRoo/frbroo_v_2.4.pdf

104. **Bergeron**, Bryan. Essentials of Knowledge Management [online]. Hoboken, New Jersey : John Wiley & Sons, Inc., 2003. [Viewed 20.06.2018]. Available from: http://www.cos.ufrj.br/~jano/LinkedDocuments/_papers/aula06/Wiley%20-20Essentials%20of%20Knowledge%20Management.pdf
105. **Bergman**, Michael K. *An Intrepid Guide to Ontologies* [online]. [Viewed 20.06.2018]. Available from: <http://www.mkbergman.com/?p=374>
106. **Berners-Lee**, T., **Fielding**, R., **Masinter**, L. *RFC 3986 URI Generic Syntax* [online]. [Viewed 21.06.2018]. Available from: <https://www.ietf.org/rfc/rfc3986.txt>
107. **Berners-Lee**, Tim, **Hendler**, James, **Lassila**, Ora. The Semantic Web: A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. In: *Scientific American* [online]. May 2001 [Viewed 28.06.2018]. Available from: https://www-sop.inria.fr/acacia/cours/essi2006/Scientific%20American_%20Feature%20Article_%20The%20Semantic%20Web_%20May%202001.pdf
108. **Berners-Lee**, Tim. *Semantic Web on XML* [online]. Washington DC 2000/12/06. Slides [Viewed 28.06.2018]. Available at: <http://www.w3.org/2000/Talks/1206-xml2k-tbl>
109. **Biagetti**, Maria Teresa. Digital libraries and semantic searching. In: *Knowledge organization in the 21th century: between historical patterns and future prospects* [online]. Proceedings of the Thirteenth International ISKO Conference, 19–22 May 2014, Kraków, Poland. Würzburg: Ergon Verlag, 2014, p. 486–492. [Viewed 16.06.2018]. Available from: https://www.researchgate.net/publication/263924292_Digital_Libraries_and_semantic_searching_2014;
110. **Bliss**, Henry. *The Organization of Knowledge and the System of the Sciences* [online]. New York: H. Holt and Company, 1929. [Viewed 20.06.2018]. Available from: <https://babel.hathitrust.org/cgi/pt?id=mdp.39015011706697;view=1up;seq=7>
111. **Blumauer**, A., **Pellegrini**, T. Semantic Web und semantische Technologien [online]. In: *Semantic Web: Wege zur vernetzten Wissensgesellschaft*. Berlin, Heidelberg: Springer. [Viewed 20.06.2018]. Available from: https://www.researchgate.net/publication/321620314_Semantic_Web_Wege_zur_vernetzten_Wissensgesellschaft
112. **Boeuf**, Patrick Le, **Doerr**, Martin, **Ore**, Christian Emil, **Stead**, Stephen, eds. *Definition of the CIDOC Conceptual Reference Model Version Version 6.2.3 May 2018* [online]. [Viewed 24.05.2018]. PDF, 58.05KB. Available from: <http://www.cidoc-crm.org/Version/version-6.2.3-0> .
113. **Bœuf**, Patrick Le, ed. *PRESSoo: Extension of CIDOC CRM and FRBRoo for the modelling of bibliographic information pertaining to continuing resources* [online]. Version 1.3. August 2016. [Viewed 24.05.2018]. PDF. Available from: https://www.ifla.org/files/assets/cataloguing/PRESSoo/pressoo_v1-3.pdf
114. **Bratková**, Eva, **Kucerová**, Helena. Knowledge Organization Systems and Their Typology [online]. In: *Revue of Librarianship*, 2014, 25, № 2, p. 8–9. [Viewed 20.06.2018]. Available from:
115. **Break** down language barriers. In: *OCLC* [online]. Dewey services. Overview. [Viewed 21.06.2018]. Available from: <https://www.oclc.org/en/dewey.html>

116. **BRICKS** Project ID: 507457. In: *CORDIS* [online] Funded under: FP6-IST. [Viewed 20.06.2018]. Available from: https://cordis.europa.eu/project/rcn/71256_en.html
117. **BRICKS** Project: *Building Resources for Integrated Cultural Knowledge Services* [online]. IST 507457 [Viewed 20.06.2018]. Available from: <http://www.brickscmmunity.org>
118. **Broughton**, Vanda, **Hansson**, Joacim, **Hjørland**, Birger, **Huertas**, María José López. Knowledge Organization [online]. In: *European Curriculum Reflections on Library and Information Science Education*. Copenhagen: Royal School of Information Science, p. 133–148. [Viewed 28.06.2018]. Available from: <http://www.webcitation.org/6q0hjT516>
119. **Broughton**, Vanda. *Essential thesaurus construction*. London: Facet Publishing, 2006.
120. **BS 8723-1:2005** Structured vocabularies for information retrieval. Guide. Definitions, symbols and abbreviations.
121. **BS 8723-2:2005** Structured vocabularies for information retrieval. Guide. Thesauri.
122. **BS 8723-3:2007** Structured vocabularies for information retrieval. Guide. Vocabularies other than thesauri.
123. **Burnett**, Kathleen, Ng, Kwong Bor, **Park**, Soyeon. A Comparison of the Two Traditions of Metadata Development. In: *Journal of the American Society for Information Science* [online], Vol. 50, 1999, Issue 13, p. 1209–1217. [Viewed 20.06.2018]. Available from: https://www.researchgate.net/publication/220433138_A_Comparison_of_the_Two_Tr aditions_of_Metadata_Development
124. **Candela**, L., **Nardi**, A. eds. D3.2b The Digital Library Reference Model [online]. [Viewed 27.09.2017]. PDF, 12.1 Mb. Available from: <http://bscw.research-infrastructure.eu/pub/bscw.cgi/d222816/D3.2b%20Digital%20Library%20Reference%20Model.pdf>
125. **CERIF**. In: *euroCRIS Current Research Information Systems*. [online]. [Viewed 24.05.2018]. Available from: <https://www.eurocris.org/cerif/main-features-cerif>
126. **CERIF-1.6**. In: *euroCRIS Current Research Information Systems* [online]. [Viewed 24.05.2018]. Available from: <http://www.eurocris.org/cerif/feature-tour/cerif-16>
127. **Chayka**, Kyle. The Library of Last Resort. In: *n+1* [online]. Online only. [Viewed 28.06.2018]. Available from: <https://nplusonemag.com/online-only/online-only/the-library-of-last-resort/>
128. **Chen**, P. P. The Entity-Relationship Model – Towards a Unified View of Data. In: *ACM Transactions on database Systems* [online]. Vol. 1, No.1, March, 1976. [Viewed 19.08.2017]. Available from: <https://pdfs.semanticscholar.org/1348/1add6a6ce710af7004577ccb0ea7f92e887d.pdf>
129. **Chu**, Heting. *Information Representation and Retrieval in the Digital Age* [online]. 2nd ed. Medford, N.J.: Information Today Inc., 2010. [Viewed 20.06.2018]. Available from: <https://www.google.com/books?hl=bg&lr=&id=Rzg6WagUrawC&oi=fnd&pg=PR13&dq=Information+Representation+and+Retrieval+in+the+Digital+Age&ots=YHGoIyhQP&sig=Yul8UyU489acvi501JI5g6W0xH0>

130. **Clarke**, Stella Dextre, **Zeng**, Marcia Lei. *From ISO 2788 to ISO 25964: the evolution of thesaurus standards towards interoperability and data modeling*. In: *Information Standards Quarterly* [online]. Winter 2012, Vol. 24, Issue 1, p. 20–26. [Viewed 05.06.2018]. Available from: https://www.researchgate.net/publication/275269295_From_ISO_2788_to_ISO_25964_the_evolution_of_thesaurus_standards_towards_interoperability_and_data_modeling
131. **Classification** Scheme Source Codes. In: *Source Codes for Vocabularies, Rules, and Schemes* [online]. [Viewed 21.06.2018}. Available from: <http://www.loc.gov/standards/sourcelist/classification.html>
132. Classification. In: *The Intellectual Foundation of Information Organization* [online]. Cambridge, Massachusetts: MIT Press, 2000, p. 126–146. [Viewed 20.06.2018]. Available from: <https://sites.evergreen.edu/wp-content/uploads/sites/226/2016/08/svenonius-chap8.pdf>
133. **Cleverdon**, Cyril W. *Aslib Cranfield research project: report on the testing and analysis of an investigation into the comparative efficiency of indexing systems* [online]. Cranfield: College of Aeronautics, 1962. [Viewed 20.06.2018]. Available from: <http://hdl.handle.net/1826/836>
134. **Cleverdon**, Cyril, **Keen**, Michael. *Aslib cranfield research project factors determining the performance of indexing systems. Vol. 2*. [online]. Cranfield, 1966.. [Viewed 20.06.2018]. Available from: http://sigir.org/files/museum/Factors_Determining_the_Performance_of_Indexing_Systems_Vol_2/pdfs/p252-chapter_8.pdf
135. **Committee** on Cataloging: Description and Access. Task Force on Metadata. Final Report [online] CC:DA/TF/Metadata/5 June 16, 2000 In: *ALCTS: Association for Library Collections & Technical Services*. [Viewed 21.06.2018]. Available from: <http://downloads.alcts.ala.org/ccda/tf-meta6.html>
136. *The Common European Research Information Format CERIF Semantic Vocabulary 0.2*. [online]. Namespace Document 01 February [2013]. Vocabulary based on CERIF Model 1.5. [Viewed 24.05.2018]. Available from: <https://www.eurocris.org/ontologies/semcerif/#sec-glance>
137. **Computational** Linguistics. In: *Stanford Encyclopedia of Philosophy*. [Viewed 28.02.2018]. Available from: <https://plato.stanford.edu/entries/computational-linguistics/>
138. **Concise UNIMARC Classification Format** [online]. (31/10/2000). Latest Revision: November 24, 2000. [Viewed 21.06.2018]. PDF, 462 KB. Available from: https://www.ifla.org/files/assets/unimarc/unimarc_format_classification_concise.pdf
139. **Dahlberg**, Ingetraut. *Grundlagen universaler Wissensordnung: Probleme und Möglichkeiten eines universalen Klassifikationssystems des Wissens* [online]. Pullach: Verlag Dokumentation. [Viewed 20.06.2018]. Available from: https://books.google.bg/books?id=LYnhOLiUPSIC&printsec=frontcover&hl=bg&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false
140. **Dalkir**, K. *Knowledge management in theory and practice* [online]. Burlington: Elsevier Butterworth-Heinemann, 2005. [Viewed 20.06.2018]. Available from: <https://content.taylorfrancis.com/books/download?dac=C2016-0-01883-5&isbn=9781136389757&format=googlePreviewPdf>

141. **DCMI Qualifiers** In: *Dublin Core Metadata Initiative* [online]. [Viewed 21.06.2018]. Available from: <http://dublincore.org/documents/2000/07/11/dcmes-qualifiers/>
142. **DCMI Specifications**. In: *Dublin Core Metadata Initiative* [online]. [Viewed 21.06.2018]. Available from: <http://dublincore.org/specifications/>
143. **DCMI Type Vocabulary**. DCMI Metadata Terms. In: *Dublin Core Metadata Initiative* [online]. [Viewed 21.06.2018]. Available from: <http://dublincore.org/documents/dcmi-type-vocabulary/>
144. **Digital Collections**. In: *British Library Labs: Experiment with our digital collections* [online]. [Viewed 28.06.2018]. Available from: <http://labs.bl.uk/Digital+Collections>
145. **Digital Collections**. In: The Library of Congress [online]. [Viewed 20.07.2017]. Available from: <https://www.loc.gov/collections/?c=160&sp=2>
146. **Digital Library Federation** [1998] [online]. [Viewed 19.08.2017]. Available from: <http://old.diglib.org/about/dldefinition.htm>
147. **Doerr**, Martin. The CIDOC CRM, an Ontological Approach to Schema Heterogeneity. In: *Semantic Interoperability and Integration. Dagstuhl Seminar Proceedings 04391* [online]. Dagstuhl: Schloss Dagstuhl, 2005. [Viewed 24.05.2018]. PDF, 0.2 MB. Available from: <https://pdfs.semanticscholar.org/c40f/1da72c8c952437135843a3947f56533c334c.pdf> .
148. **Dowlin**, Kenneth E. The Electronic Library. New York : Neal-Schuman Publishers, 1984.
149. Dublin Core Metadata Element Set, Version 1.1: Reference Description. In: *Dublin Core Metadata Initiative* [online]. [Viewed 21.06.2018]. Available from: <http://dublincore.org/documents/dces/>
150. **Early records**. In: *University of Cambridge* [online]. About the university. [Viewed 03.09.2017]. Available from: <https://www.cam.ac.uk/about-the-university/history/early-records>
151. **EMIS** [online]: *European Mathematical Information Service: Electronic Resources for Mathematicians since 1995*. [Viewed 03.09.2017]. Available from: <https://www.emis.de/index.html>
152. **Encounter** Global History. In: *Princeton University* [online]. Library. [Viewed 25.05.2018]. Available from: <https://www.princeton.edu/academics/library>
153. **Entity–relationship model** [online]. [Viewed 24.05.2018]. Available from: https://en.wikipedia.org/wiki/Entity%E2%80%93relationship_model
154. **EPrints** [online]. In: *Wikipedia*. [Viewed 24.05.2018]. Available from: <https://en.wikipedia.org/wiki/EPrints>
155. **European Guide to good Practice in Knowledge Management – Part 1: Knowledge Management Framework** [online]. CEN Workshop Agreement CWA 14924-1. Brussels: European Committee For Standardization, March 2004. [Viewed 20.06.2018]. Available from: <http://www.fraserhealth.ca/media/Euro%20Guide%20to%20good%20practice%20in%20KM%20Part%201.pdf>;
156. **Factsheet**. In: *Greenstone digital library software* [online]. [Viewed 24.05.2018]. Available from: <http://www.greenstone.org/factsheet>

157. **Fascinating** Facts. Statistics. In: *The Library of Congress* [online]. [Viewed 20.07.2017]. Available from: <https://www.loc.gov/about/fascinating-facts/>
158. **Fox**, Edward A., **Torres**, Ricardo da Silva. *Digital Library Technologies: Complex Objects, Annotation, Ontologies, Classification, Extraction, and Security* [online]. San Rafael, CA: Morgan & Claypool, 2014, p. XXV. [Viewed 20.06.2018]. Available from: https://books.google.bg/books?id=kWA-AwAAQBAJ&printsec=frontcover&dq=Digital+Library+Technologies&hl=bg&sa=X&ved=0ahUKEwjA24K-sa_bAhVGJpoKHxf9C_wQ6AEIJjAA#v=onepage&q=Digital%20Library%20Technologies&f=false
159. **Function**, collections, statistics. In: *ETH Zurich: ETH Library* [online]. [Viewed 25.05.2018]. Available from: <http://www.library.ethz.ch/en/About-us/Function-collections-statistics>
160. **Functional** requirements for subject authority data (FRSAD): A conceptual model [online]. [Viewed 17.07.2017]. PDF. Available from: <http://www.ifla.org/files/assets/classification-and-indexing/functional-requirements-for-subject-authority-data/frsad-final-report.pdf>
161. **Gärdenfors**, Peter. *How to make the Semantic Web more semantic* [online]. [Viewed 20.06.2018]. Available from: <https://slab.org/tmp/Gardenfors04.pdf>
162. **General** Information. In: *The Library of Congress* [online]. [Viewed 20.07.2017]. Available from: <https://www.loc.gov/about/general-information/>
163. **Geschichte** der Universitätsbibliothek Heidelberg. In: *Universitätsbibliothek* [online]. [Viewed 26.05.2018]. Available from: <http://www.ub.uni-heidelberg.de/allg/profil/geschichte.html>
164. **Greenstone** digital library software [online]. [Viewed 24.05.2018]. Available from: <http://www.greenstone.org/>
165. **GRL2020** Europe Call for Action Outcomes and recommendations of GRL2020 Asia [online]. PDF. [Viewed 26.05.2018]. Available from: <http://www.grl2020.net/index.html>
166. **GRL2020** Mission [online]. In: *GRL2020 A Vision for Global Research Libraries* [Viewed 26.05.2018]. Available from: <http://www.grl2020.net/index.php/about-grl2020.html>
167. **Gruber**, T. R. A Translation Approach to Portable Ontology Specifications [online]. In: *Knowledge Acquisition*, 1993, 2, 199-220 [Viewed 20.06.2018]. Available from: <http://tomgruber.org/writing/ontolingua-kaj-1993.pdf>
168. **Gruber**, T. R. Collective Knowledge Systems: Where the Social Web meets the Semantic Web. In: *Journal of Web Semantics* [online]. 2008, V. 6, № 1, p. 4–13. [Viewed 28.06.2018]. Available from: <http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=D4138B8FE67DEB6EEA61CCF8FC9E70B8?doi=10.1.1.106.183&rep=rep1&type=pdf>
169. **Guarino**, N. Formal Ontology and Information Systems. In: *Formal Ontology in Information Systems: Proceedings of the 1st International Conference* [online]. Trento, Italy, June 6–8 1998. Amsterdam: IOS Press. p. 3–15 [Viewed 24.05.2018]. PDF, 58.05KB. Available from: <https://klevas.mif.vu.lt/~donatas/Vadovavimas/Temos/OntologiskaiTeisingasKoncepti>

nisModeliavimas/papildoma/Guarino98-
Formal%20Ontology%20and%20Information%20Systems.pdf

170. **Guarino**, Nicola, **Carrara**, Massimiliano, **Giaretta**, Pierdaniele. Formalizing Ontological Commitment. In: *Proceedings of the 12th National Conference on Artificial Intelligence* [online]. Seattle, WA, USA, July 31–August 4, 1994, Vol. 1. Massachusetts: MIT Press, 1994, p. 560–567 [Viewed 24.05.2018]. Available from: <https://pdfs.semanticscholar.org/3350/b79fa347d667188b5f3d1fbbc3ac3d84fcf2.pdf>.
171. **Guarino**, Nicola, **Oberle**, Daniel, **Staab**, Steffen. What Is an Ontology? In:
172. *Handbook on Ontologies* [online]. Berlin-Heidelberg: Springer, 2009. [Viewed 20.06.2018]. Available from: https://www.researchgate.net/publication/226279556_What_Is_an_Ontology
173. *Handbook on Ontologies* [online]. Berlin-Heidelberg: Springer, 2009 [Viewed 20.06.2018]. Available from: https://www.researchgate.net/publication/226279556_What_Is_an_Ontology
174. **Harvard** Library Annual Report FY 2013. In: *Harvard University* [online]. [Viewed 25.05.2018]. Available from: <http://library.harvard.edu/annual-report-fy-2013>.
175. **Heidelberg** University Nobel Laureates. In: *Universität Heidelberg Zukunft Seit 1386* [online]. [Viewed 26.05.2018]. Available from: <https://www.uni-heidelberg.de/university/history/nobellaureates.html>
176. **Hirst**, Graeme. *Ontology and the Lexicon* [online]. [Viewed 20.06.2018]. Available from: <ftp://ftp.cs.toronto.edu/pub/gh/Hirst-Ontol-2009.pdf>
177. *The history of conceptual modeling* [online]. [Viewed 24.05.2018]. Available from: <http://cs-exhibitions.uni-klu.ac.at/index.php?id=185>
178. **History** of Heidelberg University. In: *Universität Heidelberg Zukunft Seit 1386* [online]. [Viewed 26.05.2018]. Available from: <https://www.uni-heidelberg.de/university/history/history.html>
179. **History** of the Bodleian. In: *Bodleian Library&Radcliffe Camera: University of Oxford* [online]. [Viewed 25.05.2018]. Available from: <http://www.bodleian.ox.ac.uk/bodley/about-us/history>
180. **Hjørland**, Birger. 2007. Semantics and Knowledge Organization [online]. In: *Annual Review of Information Science and Technology*, 2007, Vol. 41, Issue 1, p. 367–405 [Viewed 28.06.2018]. Available from: https://www.researchgate.net/publication/277803483_What_is_Knowledge_Organization_KO
181. **Hjørland**, Birger. Fundamentals of Knowledge Organization [online]. In: *Knowledge Organization*, 2003, Vol. 30, Issue 2, pp. 87–111. [Viewed 20.06.2018]. PDF, 1MB. Available from: <http://ppggoc.eci.ufmg.br/downloads/bibliografia/Hjorland2003.pdf>
182. **Hjørland**, Birger. Knowledge organization. *Knowledge Organization*, 2016, Vol. 43, Issue 6, pp. 475-484. Also available in Hjørland, Birger, ed. *ISKO: Encyclopedia of Knowledge Organization* [online]. [Viewed 20.06.2018]. Available from: http://www.isko.org/cyclo/knowledge_organization
183. **Hjørland**, Birger. What Is Knowledge Organization (KO) [online]. In: *Knowledge Organization*, 2008, Vol. 35, Issue 2–3), p. 86–102. [Viewed 28.06.2018]. Available from:

- https://www.researchgate.net/publication/277803483_What_is_Knowledge_Organization_KO;
184. **Hjørland**, Birger: *Automatic indexing*. [Viewed 21.06.2018]. Available from: http://www.iva.dk/bh/lifeboat_ko/CONCEPTS/automatic_indexing.htm
 185. **Hodge**, Gail. Systems of Knowledge Organization for Digital Libraries: Beyond Traditional Authority Files. In: *CLIR: Council on Library and Information Resources* [online]. April 2000. [Viewed 20.06.2018]. Available from: <http://www.clir.org/pubs/reports/pub91/contents.html>
 186. **Hutchins**, John. Retrospect and prospect in computer-based translation. In: *Proceedings of MT Summit VII* [online]. 1999, pp. 30–44. [Viewed 20.06.2018]. Available from:
 187. **IFLA Library Reference Model: A Conceptual Model for Bibliographic Information** [online], August 2017 Revised after world-wide review. [Viewed 19.08.2017]. PDF, 4.73 MB, Available from: <https://www.ifla.org/files/assets/cataloguing/frbr-lrm/ifla-lrm-august-2017.pdf>
 188. **IFLA. Guidelines for Subject Access in National Bibliographies** [online]. Draft May 2011 for Worldwide Review. IFLA Working Group on Guidelines for Subject Access by National Bibliographic Agencies [Viewed 21.06.2018]. PDF, 7.5 MB. Available from: https://www.ifla.org/files/assets/classification-and-indexing/subject-access-by-national-bibliographic-agencies/nba_guidelines_draft_2011-05.pdf
 189. **In Memoriam**. In: *Cornel CIS Computer Science* [online]. Last modified by Karla Consroe on 05/31/2018 10:48:21 [Viewed 20.06.2018]. Available from: <http://www.cs.cornell.edu/Info/Department/Annual96/Beginning/salton.html>
 190. **Interoperability** with SKOS and other schemas. In: *ISO 25964 – the international standard for thesauri and interoperability with other vocabularies* [online]. [Viewed 20.06.2018]. Available from: <https://www.niso.org/schemas/iso25964#skos>
 191. **Interoperability**. In: **Reitz**, Joan M. *ODLIS – Online Dictionary for Library and Information Science* [online]. [Viewed 24.05.2018]. Available from: http://www.abc-clio.com/ODLIS/odlis_i.aspx
 192. **Interoperability**. In: *Standards Glossary* [online]. [Viewed 10.06.2017]. Available from: https://www.ieee.org/education_careers/education/standards/standards_glossary.html .
 193. **Introducing** the Cambridge Digital Library. In: *Cambridge Digital Library* [online]. [Viewed 26.05.2018]. Available from: <https://cudl.lib.cam.ac.uk/about/>
 194. **Introduction** and history [online]. In: *University of Oxford* [Viewed 03.08.2017]. Available from: <https://www.ox.ac.uk/about/organisation/history?wssl=1>
 195. **Introduction**. In: *MARC 21 Format for Bibliographic Data* [online]. October 2006. [Viewed 20.06.2018]. Available from: <http://www.loc.gov/marc/bibliographic/bdintro.html>
 196. **Invenio**. In: *Wikipedia* [online]. [Viewed 24.05.2018]. Available from: <https://en.wikipedia.org/wiki/Invenio>
 197. **JeromeDL e-library with semantics** [online]. [Viewed 24.06.2018]. Available from: http://www.isin.ie/assets/13/1E513B27-00B2-64D5-516A8120E911BA4E_document1/1__jeromedl.pdf

198. **JULAC** Shared ILS. In: *JULAC: Joint University Librarians Advisory Committee* [online]. [Viewed 26.05.2018]. Available from: http://www.julac.org/?page_id=5499
199. **Knowledge** retrieval. In: *Wikipedia* [online]. [Viewed 20.06.2018]. Available from: https://en.wikipedia.org/wiki/Knowledge_retrieval
200. **Koch**, Traugott, **Vizine-Goetz**, Diane. Automatic classification and content navigation support of Web services: DESIRE II cooperates with OCLC [online]. Annual Review of OCLC Research.1998.
201. **Kresh**, D. *The Whole Digital Library Handbook* [online]. Chicago, IL: American Library Association Editions, 2007, 416 p. [Viewed 24.05.2018]. Available from: [Irc.tnu.edu.vn/upload/collection/brief/7805_0838909264.pdf](http://irc.tnu.edu.vn/upload/collection/brief/7805_0838909264.pdf)
202. **Kruk**, S. R., **Decker**, S., **Gzella**, A., **Grzonkowski**, S., **McDaniel**, B. Social semantic collaborative filtering for digital libraries. In: *Journal of Digital Information* [online]. Special Issue, 2006. [Viewed 24.05.2018]. Available from: https://www.researchgate.net/profile/Sebastian_Kruk2/publication/242404395_Social_Semantic_Collaborative_Filtering_for_Digital_Libraries/links/5568061908aeccd777378f9d.pdf
203. **Kruk**, S. R., **Woroniecki**, T., **Gzella**, Adam, **Dabrowski**, Maciej, **McDaniel**, Bill. *The anatomy of a Social Semantic Digital Library?* [online]. [Viewed 24.05.2018]. Available from: https://www.researchgate.net/publication/267718960_The_anatomy_of_a_Social_Semantic_Digital_Library
204. **Kruk**, Sebastian Ryszard, **Decker**, Stefan, **Zieborak**, Lech. JeromeDL – Adding Semantic Web Technologies to Digital Libraries. In: *Database and Expert Systems Applications: 16th International Conference* [online]. DEXA 2005, Copenhagen, Denmark, August 22–26, 2005, Proceedings, p.716–725. [Viewed 20.06.2018]. Available from: https://aran.library.nuigalway.ie/bitstream/handle/10379/644/70ae42e7_1.pdf?sequence=1&isAllowed=y
205. **Kruk**, Sebastian Ryszard, **Decker**, Stefan, **Zieborak**, Lech. *JeromeDL-Reconnecting Digital Libraries and the Semantic* [online]. [Viewed 20.06.2018]. Available from: https://www.researchgate.net/publication/228994759_JeromeDL-Reconnecting_Digital_Libraries_and_the_Semantic_Web
206. **Kruk**, Sebastian Ryszard, **Decker**, Stefan. *Semantic Social Collaborative Filtering with FOAFRealm* [online] [Viewed 20.06.2018]. Available from: https://aran.library.nuigalway.ie/bitstream/handle/10379/501/5e68d6a6_1.pdf?sequence=1&isAllowed=y
207. **Kruk**, Sebastian Ryszard, **Haslhofer**, Bernhard, **Piotrowski**, Piotr, **Westerski**, Adam, **Woroniecki**, Tomasz. *The Role of Ontologies in Semantic Digital Libraries* [online]. [Viewed 20.06.2018]. Available from: https://www.researchgate.net/publication/242734216_The_Role_of_Ontologies_in_Semantic_Digital_Libraries
208. **Kruk**, Sebastian Ryszard, **Kruk**, Ewelina, **Stankiewicz**, Katarzyna. *Evaluation of Semantic and Social Technologies for Digital Libraries* [online]. [Viewed 24.05.2018]. Available from: https://aran.library.nuigalway.ie/xmlui/bitstream/handle/10379/395/3debevku_1.pdf?sequence=1

209. **Kruk**, Sebastian Ryszard, **McDaniel**, Bill. Goals of Semantic Digital Libraries. In: *Semantic Digital Libraries* [online]. Berlin Heidelberg: Springer.
210. **Kruk**, Sebastian Ryszard, **Synak**, Marcin, **Zimmerman**, Kerstin. *MarcOnt – Integration Ontology for Bibliographic Description Formats* [online]. [Viewed 28.06.2018]. Available from: https://www.researchgate.net/publication/228867665_Marcont-integration_ontology_for_bibliographic_description_formats
211. **Kruk**, Sebastian Ryszard, **Woroniecki**, Tomasz, **Gzella**, Adam, **Dabrowski**, Maciej. *JeromeDL – a Semantic Digital Library* [online]. [Viewed 20.06.2018]. Available from: <https://pdfs.semanticscholar.org/1a7b/2118cbc7ff520d26a9397bbb2692be68f056.pdf>
212. **Kruk**, Sebastian Ryszard, **Woroniecki**, Tomasz, **Gzella**, Adam, **Dabrowski**, Maciej. JeromeDL: The Social Semantic Digital Library. In *Semantic Digital Libraries* [online]. Berlin Heidelberg: Springer, 2009.
213. **Kruk**, Sebastian Ryszard. FOAF-Realm – control your friends' access to the resource [online] [Viewed 28.06.2018]. Available from: https://www.w3.org/2001/sw/Europe/events/foaf-galway/papers/fp/foaf_realm/
214. **Larson**, Ray L. Experiments in automatic Library of Congress Classification In: *Journal of the American Society for Information Science*. 1992, № 43, p. 130–148; CADIAL project. [Viewed 21.06.2018]. Available from: <http://www.cadial.org/index.php?>
215. **Lesk**, M. *Practical Digital Libraries: Books, Bytes and Bucks* [online]. San Francisco: Morgan Kaufman, 1997. [Viewed 19.08.2017]. Available from: <https://books.google.bg/>
216. **Lesk**, M. *Understanding Digital Libraries*. San Francisco, CA: Morgan Kaufmann Publishers, 2005, p. 456;
217. **Libraries** Directory. In: *University of Cambridge* [online]. [Viewed 26.05.2018]. Available from: http://www.lib.cam.ac.uk/libraries_directory/libraries_directory_n.cgi
218. **Libraries**. In: *University of Oxford* [online]. [Viewed 25.05.2018]. Available from: <https://www.ox.ac.uk/research/libraries?wssl=1>
219. **Library**. In: *Princeton University* [online]. [Viewed 25.05.2018]. Available from: <https://www.princeton.edu/academics/library>
220. **Lucarelli**, Anna, **Cheti**, Alberto, **Paradisi**, Federica. Subject indexing in Italy: Recent advances and future perspectives. In: *Programme and Proceedings* [online]. 75th IFLA General Conference and Assembly. Milan, Italy, 23–27 Aug. 2009, [Viewed 21.06.2018]. Available from: <http://www.ifla.org/files/hq/papers/ifla75/200-lucarelli-en.pdf>
221. **MADS** Schema & Documentation In: *MODS: Metadata Object Description Schema* [online]. [Viewed 20.06.2018]. Available from: <http://www.loc.gov/standards/mads/>
222. **Mai**, Chan Lois. Library of Congress Classification in a New Setting: Beyond Shelfmarks. In: *Library of Congress* [online]. Classification Distribution Service. [Viewed 21.06.2018]. Available from: <http://2008.myvote.org/www.loc.gov/cds/chanarticle.html>

223. **Mann**, Thomas. Will Google's keyword searching eliminate the need for LC cataloguing and classification?. In: *Internet Archive* [online]. Last updated on August 15, 2005. [Viewed 20.06.2018]. Available from: <http://www.guild2910.org/searching.htm>
224. **MARC 21 Format for Bibliographic Data** [online]. October 2006. [Viewed 20.06.2018]. Available from: <http://www.loc.gov/marc/bibliographic/bdintro.html>
225. **MARC 21 Format for Classification Data** [online]. 2000 Edition Update No. 1 (October 2000) through Update No. 26 (April 2018). [Viewed 21.06.2018]. Available from: <https://www.loc.gov/marc/classification/>
226. **MARC 21 формат за библиографски данни**. Актуализация № 1 (октомври 2001 г.) – актуализация № 13 (септември 2011 г.). Достъпно от: <http://www.nalis.bg/MARC-BIB/>;
227. **MARC 21 формат за библиографски данни: Кратка версия** [online]. 2010. [Прегледан 20.06.2018]. Достъпно от: <http://www.nalis.bg/MARC-BIB/concise/bdleader.html>
228. **MARC to Dublin Core Crosswalk** [online]. Development and MARC Standards Office
229. **MARC21 Lite Формат за машинночетими библиографски записи**. София: Унив. изд. „Св. Климент Охридски“, 2005.
230. **Marchionini, G., Fox, E.**: Progress toward digital libraries: Augmentation through integration. *Information Processing & Management*, 35(3), 1999, p. 219–225.
231. **MARCXML: MARC 21 XML Schema** [online]. [Viewed 20.06.2018]. Available from: <http://www.loc.gov/standards/marcxml/>
232. **Markey, Karen**. Forty Years of Classification Online: Final Chapter or Future Unlimited?. In: *Cataloguing and Classification Quarterly*, 2006. Vol. 42, № 3, p. 1–63.
233. **Massachusetts Institute of Technology Libraries** [online]. In: *Wikipedia* [Viewed 03.08.2017]. Available from: https://en.wikipedia.org/wiki/Massachusetts_Institute_of_Technology_Libraries
234. **McGray, A., Gallagher, M.** Principles for Digital Library Development. In: *Communications of the ACM* [online]. 44(5), p. 49–54, 2001 [Viewed 15.04.2018]. Available from: https://www.researchgate.net/publication/220422376_Principles_for_Digital_Library_Development
235. **Miller, George A., Beckwith, Richard, Fellbaum, Christiane, Gross, Derek, Miller, Katherine**. *Introduction to WordNet: An On-line Lexical Database* [online]. Revised August 1993. [Viewed 28.02.2018]. Available from: <http://wordnetcode.princeton.edu/5papers.pdf>
236. **Minsky, Marvin**. *A framework for representing knowledge* [online]. Boston, Massachusetts: MIT AI Laboratory Memo 306, June, 1974. [Viewed 20.06.2018]. Available from: <http://web.media.mit.edu/~minsky/papers/Frames/frames.html>
237. **Mission & Objectives**. In: *Harvard Library* [online]. [Viewed 25.05.2018]. Available from: <https://library.harvard.edu/objectives-priorities>.

238. **MIT** Highlights Timeline [online]. In: *MIT story* [Viewed 25.05.2018]. Available from: <http://mitstory.mit.edu/mit-highlights-timeline/#event-first-wind-tunnel>
239. **MIT** Libraries. In: *MIT Massachusetts Institute of Technology* [online]. [Viewed 25.05.2018]. Available from: <http://web.mit.edu/facts/libraries.html>.
240. **Mizoguchi, R., Sinita, K., Ikeda, M.** *Knowledge Engineering of Educational Systems for Authoring System Design: A preliminary results of task ontology design* [online]. [Viewed 20.06.2018]. Available from: <http://www.ei.sanken.osaka-u.ac.jp/pub/miz/miz-eaied96.html>
241. **Moeller, G.** et al. Automatic classification of the World-Wide-Web using UDC. In: *Proceedings of the GfKI*. Heidelberg : Springer, 1999
242. **National** digital library program. In: *The Library of Congress* [online]. [Viewed 16.04.2017]. Available from: <http://memory.loc.gov/ammem/dli2/html/lcndlp.html>
243. *The National Information Infrastructure: Agenda For Action Information Infrastructure* [online]. Washington, D.C.: Executive Office of the President, 1993. [Viewed 16.04.2017]. Available from: https://archive.org/stream/04Kahle000911/04Kahle000911_djvu.txt
244. The **Nation's** Largest Libraries: A Listing By Volumes Held. In: *ALA American Library Association* [online]. [Viewed 25.05.2018]. Available from: <http://www.ala.org/tools/libfactsheets/alalibraryfactsheet22>.
245. **NISO**. Understanding metadata what is metadata, and what is it for?. National Information Standards Organization (NISO), 2017, p. 7. https://groups.niso.org/apps/group_public/download.php/17446/Understanding%20Metadata.pdf
246. **Olensky, Marlies.** Semantic interoperability in Europeana. An examination of CIDOC CRM in digital cultural heritage documentation. In: *Bulletin of IEEE Technical Committee on Digital Libraries* [online]. 2010, Vol. 6, Issue 2 [Viewed 30.05.2018]. Available from: <http://www.ieee-tcdl.org/Bulletin/v6n2/Olensky/olensky.html#18>
247. **Oliver, Chris.** *Identifying resources: FRBR and accessibility* [online]. [Viewed 24.05.2018]. Available from: <http://library.ifla.org/1368/1/200-oliver-en.pdf>
248. **Olson, Hope A., Boll, John J.** *Subject analysis in online catalogs* [online]. Englewood, Colorado: Libraries Unlimited, 2001. [Viewed 20.06.2018]. Available from: https://archive.org/details/subjectanalysisi00alu_34g
249. **Our** History. In: *Princeton University* [online]. [Viewed 25.05.2018]. Available from: <https://www.princeton.edu/meet-princeton/history>
250. **Outline** of Elements and Attributes in MODS Version 3.5. In: *MODS: Metadata Object Description Schema* [online]. [Viewed 20.06.2018]. Available from: <http://www.loc.gov/standards/mods/mods-outline-3-5.html>
251. **Overview** of Text Retrieval [online]. [Viewed 20.06.2018]. Available from: <http://sifaka.cs.uiuc.edu/course/410/note/tr.html>
252. **Palomäki, Jari, Kangassalo, Hannu.** *That IS-IN Isn't IS-A: A Further Analysis of Taxonomic Links in Conceptual Modelling* [online], p. 1 [Viewed 20.06.2018]. PDF, 316 KB. Available from: http://cdn.intechopen.com/pdfs/36654/InTech-That_is_in_isn_t_is_a_a_further_analysis_of_taxonomic_links_in_conceptual_modelling.pdf

253. **Pandey**, Shriram, **Panda**, K. C. Semantic solutions for the digital libraries based on semantic web technologies. In: *Annals of Library and Information Studies* [online]. Vol. 61, December 2014, pp. 286–293 [Viewed 26.06.2018]. Available from: <https://pdfs.semanticscholar.org/2a90/cd95c53bfee550da17f5fbb6139f4cea8685.pdf>
254. **Patel**, M., **Koch**, T., **Doerr**, M., and **Tsinaraki**, C. *DELOS: A Network of Excellence in Digital Libraries* [online]. D5.3.1: Semantic Interoperability in Digital Library Systems. Task 3: Semantic Interoperability WP5: Knowledge Extraction and Semantic Interoperabilityq July 2004–June 2005. [Viewed 30.05.2018]. PDF. Available from: <http://delos-wp5.ukoln.ac.uk/project-outcomes/SI-in-DLs/SI-in-DLs.pdf>.
255. **Patel**, Manjula, **Koch**, Traugott, **Doerr**, Martin, **Tsinaraki**, Chrisa. Semantic Interoperability in Digital Library Systems [presentation]. In: *DELOS WP5 Workshop: Semantic Interoperability in DL systems* [online]. 17th September 2004, Bath, UK. [Viewed 24.05.2018]. Available from: http://opus.bath.ac.uk/35098/4/WP5_Workshop_040916.pdf.
256. **Portrait** 2018: „Where the future begins?”. In: *ETH Zurich* [online]. [Viewed 25.05.2018]. Available from: <https://www.ethz.ch/en/the-eth-zurich/information-material/information-brochure.html>
257. **Riva**, Pat, **Bœuf**, Patrick Le, **Žumer**, Maja, eds. *IFLA Library Reference Model: A Conceptual Model for Bibliographic Information* [online]. August 2017 Revised after world-wide review. [Viewed 19.08.2017]. PDF, 4.73 MB. Available from: <https://www.ifla.org/files/assets/cataloguing/frbr-lrm/ifla-lrm-august-2017.pdf>
258. **Rowley**, Jennifer. The controlled versus natural indexing languages debate revisited: A perspective on information and retrieval practice and research. In: *Journal of Information Science*, 1994, Vol. 20, № 2. p. 108–119.
259. **Rust**, Godfrey. Metadata: The Right Approach. In: *D-Lib Magazine* [online]. July/August 1998. [Viewed 21.06.2018]. Available from: <http://www.dlib.org/dlib/july98/rust/07rust.html>
260. **Semantic** Digital Library. In: *Dictionary information science and technology* [online]. [Viewed 24.05.2018]. Available from: <https://www.igi-global.com/dictionary/?p=Semantic+Digital+Library>
261. *Semantic Web Technologies for Digital Libraries JeromeDL – A Case Study* [online]. [Viewed 24.05.2018]. Available from: <https://sites.google.com/a/isim.net.in/nikesh/semantic-digital-libraries>
262. The **Semantic** Web Vision. In: *A semantic web primer* [online]. 3rd ed. Cambridge: MIT Press, 2012, p. 1–3. [Viewed 28.06.2018]. Available from: http://prof.mau.ac.ir/images/Uploaded_files/A%20Semantic%20Web%20Primer-The%20MIT%20Press%20.PDF
263. **Shields**, Ginger. *LI-842 Automatic Indexing Assignment*. [Viewed 21.06.2018]. 2005. Available from: http://www.shieldsnetwork.com/LI842_Shields_Automatic_Indexing.pdf
264. **Shiri**, Ali, **Chase-Kruszewski**, Sarah. Knowledge organisation systems in North American digital library collections. Abstract. In: *Program*, 2009, Vol. 43, Issue 2, pp.121–139. [Viewed 21.06.2018]. Available from: <https://www.emeraldinsight.com/doi/pdfplus/10.1108/00330330910954352>

265. **Sicilia**, Miguel-Ángel. *The Common European Research Information Format Ontology CERIF Ontology 0.2* [online]. Namespace Document 01 February 2013. Vocabulary based on CERIF Model 1.3 [Viewed 24.05.2018]. Available from: <http://eurocris.org/ontologies/cerif/1.3/index.html>
266. **SIMILE Overview: Semantic Interoperability of Metadata and Information in unLike Environments** [online]. [Viewed 20.06.2018]. Available from: <http://web.mit.edu/dspace-dev/www/simile/resources/overview.html>
267. **Singhal**, Amit. *Modern Information Retrieval: A Brief Overview* [online]. [Viewed 20.06.2018]. Available from: <http://singhal.info/ieee2001.pdf>
268. **SKOS Simple Knowledge Organization System** [online]. [Viewed 28.06.2018]. Available from: <https://www.w3.org/2004/02/skos/>
269. **SKOS Simple Knowledge Organization System**. [online]. ReferenceW3C Recommendation 18 August 2009. [Viewed 20.06.2018]. Available from: <https://www.w3.org/TR/skos-reference/>; **SKOS Simple Knowledge Organization System Reference** [online]. W3C Working Draft 25 January 2008. [Viewed 20.06.2018]. Available from: <http://www.w3.org/TR/2008/WD-skos-reference-20080125/>
270. **Soergel**, Dagobert. Digital Libraries and Knowledge Organization [online]. In: **Kruk**, Sebastian Ryszard, McDaniel, Bill, eds. *Semantic Digital Libraries*. Berlin: Springer, p. 9–39. [Viewed 20.06.2018]. Available from: https://link.springer.com/chapter/10.1007/978-3-540-85434-0_2
271. **Soergel**, Dagobert. Evaluation of Knowledge Organization Systems (KOS): Characteristics for Describing and Evaluating KOS [online]. In: *Paper presented at Classification Crosswalks: Bringing Communities Together. The 4th NKOS Workshop at ACM-IEEE Joint Conference on Digital Libraries (JCDL), June 28th, Roanoke, VA* [Viewed 20.06.2018]. Available from: <http://nkos.slis.kent.edu/2001/SoergelCharacteristicsOfKOS.pdf>
272. **Soergel**, Dagobert. *Knowledge Organization Systems: Overview* [online]. p. 3 [Viewed 20.06.2018]. Available from: <http://www.dsoergel.com/SoergelKOSOverview.pdf>
273. **Souza**, Renato Rocha, **Tudhope**, Douglas, **Almeida**, Mauricio B. Towards a Taxonomy of KOS: Dimensions for Classifying Knowledge Organization Systems [online]. In: *Knowledge Organization, Vol. 39, Issue 3*, p. 179–192. [Viewed 20.06.2018]. Available from: http://mba.eci.ufmg.br/downloads/Souza_Tudhope_Almeida_-_KOS_Taxonomy.Submitted.pdf
274. **Sowa**, John. *Knowledge representation: Logical, philosophical, and computational foundations* [online]. [Viewed 28.06.2018]. Available from: https://www.researchgate.net/publication/225070439_Knowledge_Representation_Logical_Philosophical_and_Computational_Foundations
275. **Svenonius**, Elaine. LCSH: Semantics, Syntax and Specificity. In: *Cataloging & Classification Quarterly* [online]. 2000, Vol. 29, № 1–2, p. 17–30. [Viewed 20.06.2018]. Available from: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.607.4757&rep=rep1&type=pdf>

276. **Svenonius**, Elaine. Precoordination or not?. In: *Subject indexing: principles and practices in the 90's: Proceedings of the IFLA Satellite meeting held in Lisbon, Portugal, 17–18 Aug. 1993*. München: Saur 1995, p. 231–255.
277. **Svenonius**, Elaine. Subject Languages: Introduction, Vocabulary Selection, and Classification. In: *The Intellectual Foundation of Information Organization* [online]. Cambridge, Massachusetts: MIT Press, 2000, p. 126–146. [Viewed 20.06.2018]. Available from: <https://sites.evergreen.edu/wp-content/uploads/sites/226/2016/08/svenonius-chap8.pdf>
278. **Svenonius**, Elaine. *The Intellectual Foundation of Information Organization*. Cambridge, Mass.: MIT Press, 2000.
279. **Svensson**, Lars G. National Libraries and the Semantic Web: Requirements and Applications. In: **Prasad**, A.R.D, **Madalli**, Devika P. Eds. *International Conference on Semantic Web and Digital Libraries* [online]. 2007, pp. 101–108. [Viewed 24.05.2018]. Available from: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.113.2756&rep=rep1&type=pdf>
280. **Svensson**, Lars G., **Jahns**, Yvonne. PDF, CSV, RSS and other Acronyms: Redefining the Bibliographic Services in the German National Library. In: *Programme and Proceedings* [online]. 76th IFLA General Conference and ssembly, Gothenburg, Sweden, 10–15 Aug. 2010. [Viewed 21.06.2018]. Available from: <https://www.ifla.org/past-wlic/2010/91-svensson-en.pdf>
281. **Synak**, Marcin, **Kruk**, Sebastian Ryszard. *MarcOnt Initiative the Ontology for the Librarian World* [online] [Viewed 24.06.2018]. Available from: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.114.3986&rep=rep1&type=pdf>
282. **TARDIS** *University of Southampton e-Prints* [online]. [Viewed 24.05.2018]. Available from: <http://tardis.eprints.org/>
283. **Taube**, Mortimer. Studies in coordinate indexing. In: *HathiTrust's digital library* [online]. Washington: Documentation Incorporated, 1953, 110 p. [Viewed 20.06.2018]. Available from: <https://catalog.hathitrust.org/Record/001163200>
284. **Tella**, Adeyinka, **Okojie**, Victoria, **Olaniyi**, O. T. *Social Bookmarking Tools and Digital Libraries. Key Terms in this Chapter* [online]. [Viewed 24.05.2018]. Available from: <https://www.igi-global.com/chapter/social-bookmarking-tools-and-digital-libraries/188559>
285. **Trivedi**, Mayank. Digital Libraries: Functionality, Usability, and Accessibility. In: *Library Philosophy and Practice* [online]. May 2010, p. 2. [Viewed 24.05.2018]. Available from: <https://www.webpages.uidaho.edu/~mbolin/trivedi-diglib.pdf>
286. **Tsichritzis**, Dionysios C., **Lochovsky**, Frederick H. *Data Models* [online]. New Jersey: Prentice Hall, 1982. 382 p. [Viewed 24.05.2018]. Available from: <https://books.google.bg/>
287. **Tudhope**, Douglas, **Koch**, Traugott, **Heery**, Rachel. *Terminology Services and Technology: JISC State of the Art Review* [online]. Bath: UKLON [Viewed 20.06.2018]. Available from: <http://www.ukoln.ac.uk/terminology/JISC-review2006.html>

288. **UDC** Master Reference File (MRF). In: *UDC Consortium* [online]. [Viewed 21.06.2018]. Available from: <http://www.udcc.org/index.php/site/page?view=mrf>
289. **UDC** Users Worldwide. In: *UDC consortium* [online]. [Viewed 21.06.2018]. Available from: http://www.udcc.org/index.php/site/page?view=users_worldwide
290. **United** Nations Group of Experts on Geographical Names. *Manual for the national standardization of geographical names* [online] [Viewed 20.06.2018]. Available from: New York: United Nations, 2006, p. 73-76. [Viewed 30.05.2018]. Available from: https://unstats.un.org/unsd/publication/seriesm/seriesm_88e.pdf
291. The **University's** mission and core values. In: *University of Cambridge* [online]. [Viewed 26.05.2018]. Available from: <http://www.cam.ac.uk/about-the-university/how-the-university-and-colleges-work/the-universitys-mission-and-core-values>
292. **Uschold**, Mike, **Gruninger**, Michael. Ontologies Principles, Methods and Applications. In: *KnowledgeEngineering Review* [online]. Vol. 11, № 2, June 1996 [Viewed 20.06.2018]. Available from: <http://www.aii.ed.ac.uk/publications/documents/1996/96-ker-intro-ontologies.pdf>
293. **van Heijst**, G., **Schreiber**, G. *Wielinga: Using Explicit Ontologies in Knowledge-Based System Development* *Explicit Ontologies in Knowledge-Based System Development* [online]. [Viewed 20.06.2018]. Available from: https://www.researchgate.net/publication/243779867_Wielinga_Using_Explicit_Ontologies_in_Knowledge-Based_System_Development
294. **Wang**, Georgette. Many Paths to Many Destinations. In: **Wang**, Georgette ed. *Treading different paths: informatization in Asian nations*. Norwood, N.J.: Ablex Pub. Corp., 1994. p. 247–258. ISBN 1-567500-47-1
295. **Whitley**, Richard R. *The Intellectual and Social Organization of the Sciences* [online]. Oxford: Oxford University Press, 1984 [Viewed 28.06.2018]. Available from: https://www.researchgate.net/publication/235413219_The_Intellectual_and_Social_Organization_of_the_Sciences
296. **Who** is using UDC? In: *UDC consortium* [online]. About Universal Decimal Classification (UDC). [Viewed 21.06.2018]. Available from: <http://www.udcc.org/index.php/site/page?view=about>
297. **Williams**, Robert V. *Bibliography of the history of information science and technology, 1900–2012* [online]. 8th Ed. [Viewed 15.04.2018]. Available from: <http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=ACF8FFD796125F300BD764CFB616F1F1?doi=10.1.1.676.6682&rep=rep1&type=pdf>
298. **Witten**, I. H., **Bainbridge**, D., **Nichols**, D. M. *How to Build a Digital Library*. Second Edition. Burlington, MA: Morgan Kaufmann Publishers, 2009;
299. **wnstats(7WN)**. In: *WordNet: A Lexical Database for English* [online]. [Viewed 20.06.2018]. Available from: <https://wordnet.princeton.edu/documentation/wnstats7wn>
300. **Wordnet**. In: *Секция по компютърна лингвистика* [онлайн]. [Прегледан 20.06.2018]. Достъпно от: <http://dcl.bas.bg/resursi/wordnet/>
301. **Wright**, Sue Ellen. Typology for KRRs [online]. [presentation]. In: *Networked Knowledge Organization Systems/Services/Structures NKOS*. [Viewed 20.06.2018]. Available from: <http://nkos.slis.kent.edu/2008workshop/SueEllenWright.pdf>

302. **Yi**, Kwan. *Challenges in automatic classification using library classification schemes*. [Viewed 21.06.2018]. Available from: <http://archive.ifla.org/IV/ifla72/papers/097-Yi-en.pdf>
303. **Yiyu Yao, Yi Zeng, Ning Zhong, Xiangji Huang**. Knowledge Retrieval (KR). In: *Proceedings of the 2007 IEEE/WIC/ACM International Conference on Web Intelligence* [online]. IEEE Computer Society, Silicon Valley, USA, Nov. 2–5, 2007, p. 729–735. [Viewed 15.04.2018]. Available from: https://www.researchgate.net/publication/4309759_Knowledge_Retrieval_KR
304. **Zeng**, Marcia Lei. Knowledge Organization Systems (KOS) [online]. In: *Knowledge Organization*, 2008, Vol. 35, Issue 2–3, p. 160–182 [Viewed 20.06.2018]. Available from: https://www.researchgate.net/profile/Marcia_Zeng/publication/297530633_Knowledge_organization_systems_KOS/links/59dbbaf4a6fdcc1ec89fa0fb/Knowledge-organization-systems-KOS.pdf

СПИСЪК НА ТАБЛИЦИ, ФИГУРИ И ИЛЮСТРАЦИИ

ТАБЛИЦА 1. ЙЕРАРХИЯ НА КЛАСОВЕТЕ СЪЩНОСТИ (CLASS HIERARCHY) В CIDOC CRM _____	40
ТАБЛИЦА 2. IFLA LRM ПОТРЕБИТЕЛСКИ ЗАДАЧИ. _____	46
ТАБЛИЦА 3. ИЗПОЛЗВАНИ ПРОГРАМНИ ИНСТРУМЕНТИ ЗА ЕЛЕКТРОННИ БИБЛИОТЕКИ (ЕБ) С ОТВОРЕН ДОСТЪП (OAI-PMH) И РАЗПРОСТРАНЕНИЕТО ИМ _____	57
ТАБЛИЦА 4. СИСТЕМИ ЗА ОРГАНИЗАЦИЯ НА ЗНАНИЯТА ПО ХОДЖ _____	84
ТАБЛИЦА 5. СИСТЕМИ ЗА ОРГАНИЗАЦИЯ НА ЗНАНИЯТА В SKOS. _____	113
ТАБЛИЦА 6. ТИПОВЕ МЕТАДАНИИ. _____	147
ТАБЛИЦА 7. СТРУКТУРА И ЕЛЕМЕНТИ ОТ ДАНИИ НА БИБЛИОГРАФСКОТО ОПИСАНИЕ. _	149
ТАБЛИЦА 8. ЙЕРАРХИЧНИ КЛАСИФИКАЦИИ СПОРЕД НИВО НА УНИВЕРСАЛНОСТ. _____	153
ТАБЛИЦА 9. ЛЕКСИКАЛНИ ЕДИНИЦИ В ЙЕРАРХИЧНИ КЛАСИФИКАЦИИ. _____	154
ТАБЛИЦА 10. ИНФОРМАЦИОННИ ЕЗИЦИ: ВИДОВЕ И ФУНКЦИИ. _____	164
ТАБЛИЦА 11. БЪЛГАРСКИ ЕЛЕКТРОННИ БИБЛИОТЕКИ. _____	168
ТАБЛИЦА 12. ИНФОРМАЦИОННИ ЕЗИЦИ В БЪЛГАРСКИТЕ ЕЛЕКТРОННИ БИБЛИОТЕКИ. _	179
ТАБЛИЦА 13. DCMES ЕЛЕМЕНТИ И ОПРЕДЕЛИТЕЛИ _____	194
ФИГУРА 1. РЕФЕРЕНТНА АРХИТЕКТУРА НА ЕЛЕКТРОННА БИБЛИОТЕКА _____	30
ФИГУРА 2. КОНЦЕПТУАЛНА КАРТА НА ОБЛАСТТА НА РЕСУРСА _____	33
ФИГУРА 3. КОНЦЕПТУАЛНА КАРТА НА ОБЛАСТТА НА СЪДЪРЖАНИЕТО _____	34
ФИГУРА 4. КОНЦЕПТУАЛНА КАРТА НА ОБЛАСТТА НА ПОТРЕБИТЕЛЯ _____	36
ФИГУРА 5. КОНЦЕПТУАЛНА КАРТА НА ОБЛАСТТА НА ФУНКЦИИТЕ _____	37
ФИГУРА 6. ЕЛЕМЕНТИ НА IFLA LRM _____	47
ФИГУРА 7. ОБЕКТИ И ВРЪЗКИ В IFLA LRM _____	49
ФИГУРА 8. ОБЕКТИ И ИВРЪЗКИ В CERIF _____	51
ФИГУРА 9. CERIF ВЗАИМОДЕЙСТВИЕ С ОНТОЛОГИИ _____	52
ФИГУРА 10. ОНТОЛОГИЯ НА ИНФОРМАЦИЯ ЗА НАУЧНИТЕ ИЗСЛЕДВАНИЯ CERIF. _____	53
ФИГУРА 11. ТИПОВЕ СИСТЕМИ ЗА ОРГАНИЗАЦИЯ НА ЗНАНИЯТА _____	82
ФИГУРА 12. СКАЛА НА СЕМАНТИКАТА НА СИСТЕМИТЕ ЗА ОРГАНИЗАЦИЯ НА ЗНАНИЯТА	83
ФИГУРА 13. ВИЗИЯ НА ТИМ БЪРНАРС-ЛИЙ ЗА СЕМАНТИЧНИЯ УЕБ. _____	92

ФИГУРА 14. ЗАВИСИМОСТ МЕЖДУ ПРИЛОЖЕНИЕ И УДОБСТВО НА РАБОТА НА ТИПОВЕ ОНТОЛОГИИ _____	101
ФИГУРА 15. MARCONTOLOGY В JEROMEDL _____	105
ФИГУРА 16. SKOS ОПИСАНИЕ. _____	113
ФИГУРА 17. СЕМАНТИЧНА ЕЛЕКТРОННА БИБЛИОТЕКА _____	116
ФИГУРА 18. ИНСТРУМЕНТИ НА УЕБ 2.0 _____	117
ФИГУРА 19. ПРИМЕР ОТ JEROMEDL _____	119
ФИГУРА 20. БЪЛГАРСКИ ЕЛЕКТРОННИ БИБЛИОТЕКИ СПОРЕД ПРИНАДЛЕЖНОСТ _____	175
ФИГУРА 21. СХЕМА НА МОДЕЛ НА ЛИНГВИСТИЧНО ОСИГУРЯВАНЕ НА БЪЛГАРСКА ЕЛЕКТРОННА БИБЛИОТЕКА С УНИВЕРСАЛНА ТЕМАТИКА _____	184
ИЛЮСТРАЦИЯ 1. ДИГИТАЛНА БИБЛИОТЕКА НА НАЦИОНАЛЕН СТАТИСТИЧЕСКИ ИНСТИТУТ. _____	176
ИЛЮСТРАЦИЯ 2. ТЕМАТИЧНИ РУБРИКИ НА ЕЛЕКТРОННА БИБЛИОТЕКА „МОЯТА БИБЛИОТЕКА”. _____	177
ИЛЮСТРАЦИЯ 3. ДИГИТАЛНА БИБЛИОТЕКА УНИВЕРСИТЕТ ЗА НАЦИОНАЛНО И СВЕТОВНО СТОПАНСТВО. _____	178
ИЛЮСТРАЦИЯ 4. БИБЛИОГРАФСКО ОПИСАНИЕ ОТ ЕЛЕКТРОННА БИБЛИОТЕКА „МОЯТА БИБЛИОТЕКА”. _____	178
ИЛЮСТРАЦИЯ 5. ФОРМАТ ЗА ПОКАЗВАНЕ ISBD В ДИГИТАЛНА БИБЛИОТЕКА „МОСТ ПРЕЗ ВЕКОВЕТЕ” – ТЪРГОВИЩЕ. _____	179
ИЛЮСТРАЦИЯ 6. UDC SUMMARY. _____	203
ИЛЮСТРАЦИЯ 7. ПРИМЕР ЗА НОРМАТИВЕН ЗАПИС НА ПРЕДМЕТНА РУБРИКА ОТ НОРМАТИВНАТА БАЗА ОТ ДАННИ НА НАЦИОНАЛНАТА БИБЛИОТЕКА НА ЧЕХИЯ. _	205
ИЛЮСТРАЦИЯ 8. ПРИМЕР ЗА ИНДЕКСИРАНИ ОТ СЛОВЕНСКАТА НАЦИОНАЛНА БИБЛИОТЕКА. _____	206

ПРИЛОЖЕНИЯ

Приложение 1. Йерархия на обектите в IFLA LRM: определения и примери

Йерархия на обектите		
Ниво 1 (клас)	Ниво 2 (подклас)	Ниво 3 (подклас)
<p>LRM-E1 Нещо (Res)</p> <p><i>Определение:</i> всеки обект в библиографския универсум.</p> <p><i>Примери:</i></p> <p>Париж [столица на Франция];</p> <p>Атлантида [легендарен остров];</p> <p>Агата Кристи [английска писателка];</p> <p>Одисея на Омир [древногръцко произведение].</p>		
--	<p>LRM-E2 Произведение (Work)</p> <p><i>Определение:</i> интелектуално или художествено съдържание на отделно творение.</p> <p><i>Примери:</i></p> <p>„Илюзия с огледала“ [роман на Агата Кристи];</p> <p>„Одисея“ [древногръцка епична поема на Омир];</p> <p>„Ласи се завръща“ [филм от 1943];</p> <p>„I want to hold your hand“ [песен на Джон Ленън и Пол Макартни].</p>	
--	<p>LRM-E3 Проявление (Expression)</p> <p><i>Определение:</i> разграничима комбинация от знаци, която предава интелектуално или художествено съдържание.</p> <p><i>Примери:</i></p>	

	<p>Превод от старогръцки на български език от Георги Батаклиев на „Одисея“ от Омир, 1971;</p> <p>Текст на английски език на произведението на Агата Кристи „They do it with mirrors“, original copyright 1952 [публикуван понякога и под заглавието „Murder with mirrors“];</p> <p>Dewey Decimal Classification, 23rd edition (DDC23) [английско издание];</p> <p>Classification décimale de Dewey, 23e édition [превод на френски от DDC 23 издание].</p>	
--	<p>LRM-E4 Въплъщение (Manifestation)</p> <p><i>Определение:</i> съвкупността от всички екземпляри, които имат едни и същи характеристики като интелектуално или художествено съдържание и физическа форма.</p> <p><i>Примери:</i></p> <p>Одисея / Омир ; Прев. от старогр. Георги Батаклиев с предговор от Богдан Богданов, издадена от Народна култура, 1971, в серия Световна класика;</p> <p>They do it with mirrors / Agatha Christie published in the UK by William Collins & Sons in 1952 [въплъщение на произведението];</p> <p>Murder with mirrors / Agatha Christie published in the US by Dodd, Mead & Co. in 1952 [друго въплъщение, публикувано в друга държава от друго издателство и под различно заглавие].</p>	
--	<p>LRM-E5 Единица (Item)</p> <p><i>Определение:</i> конкретен обект (или обекти) с интелектуално или художествено съдържание. От гледна точка на интелектуалното съдържание и физическата форма, единица, която е образец от въплъщение, обикновено е идентична. В много случаи това е един физически предмет (напр. екземпляр от еднотомна монография, една аудио касета и др.). Понякога обаче обектът, определен като</p>	

	<p>единица, обхваща повече от един физически предмет (напр. монография, издадена в два отделно подвързани тома, звукозапис на три отделни компактдиска и др.).</p> <p><i>Примери:</i></p> <p>Library of Congress Copy 2</p> <p>The Odyssey / Homer ; translated by Robert Fagles, Deluxe edition published in New York by Penguin Books in 1997, ISBN 0-670-82162-4;</p> <p>Лично копие от филма „Специалист по всичко“ издаден на 1 DVD от Националния филмов център, София, 2006, Регистрационен номер МК 50622-3/27.02.2006;</p> <p>Номериран екземпляр № 10 от библиофилско фототипно издание на „Абагар“ 1651 на Филип Станиславов, София, Ciela, 2011.</p>	
--	<p>LRM-E6 Агент (Agent)</p> <p><i>Определение:</i> обект, който може да извършва съзнателни действия, да получи права и да бъде държан отговорен за действията си.</p> <p><i>Примери:</i></p> <p>Омир;</p> <p>Българска национална телевизия;</p> <p>Симпозиум по глаукома;</p> <p>Бийтълс.</p>	
--	--	<p>LRM-E7 Лице (Person)</p> <p><i>Определение:</i> индивидуално човешко същество.</p> <p><i>Примери:</i></p> <p>Омир;</p> <p>Лицето, посочено чрез: истинско име „Charles Lutwidge Dodgson“</p>

		<p>[преподавател по математика] и псевдоним „Lewis Carroll“ [писател]; Агата Кристи; Марко Поло.</p>
--	--	<p>LRM-E8 Колектив (Collective Agent) <i>Определение:</i> организация или група индивиди и/или организации, идентифицирани с отделно наименование, действащи като едно цяло. <i>Примери:</i> International Federation of Library Associations and Institutions; Национална библиотека „Св. св. Кирил и Методий“; Детски радиофор БНР; XXVII национална конференция „Библиотеките – национална идентичност и многообразие“, София, 8-9 юни 2017; „Братя Мормарев“ колективен литературен</p>

		псевдоним на авторите на сценарии за популярни български филми – Марко Стойчев и Мориц Йомтов.
--	<p>LRM-E9 Име (Nomen)</p> <p><i>Определение:</i> название на обект, т.е. комбинация от знаци, която се използва, за да се отдели всеки обект в библиографската вселена. Всеки обект, посочен във вселената на дискурса, е избран най-малко с една име.</p> <p><i>Примери:</i></p> <p><i>Имена на лице:</i></p> <p>„Agatha Christie” начин на позоваване за лицето Dame Agatha Christie, Lady Mallowan;</p> <p>„Agatha Mary Clarissa Miller” начин на позоваване за лицето Dame Agatha Christie, Lady Mallowan;</p> <p>„Lady Mallowan” начин на позоваване за лицето Dame Agatha Christie, Lady Mallowan;</p> <p>„Mary Westmacott” начин на позоваване за лицето Dame Agatha Christie, Lady Mallowan;</p> <p>„Westmacott, Mary, 1890-1976” начин на позоваване за лицето Dame Agatha Christie, Lady Mallowan;</p> <p>„Christie, Agatha, 1890-1976” начин на позоваване за лицето Dame Agatha Christie, Lady Mallowan [предпочитана точка за достъп според RDA за нейните детективски романи и разкази];</p> <p>„Westmacott, Mary, 1890-1976” начин на позоваване за лицето Dame Agatha Christie, Lady Mallowan [предпочитана точка за достъп според RDA за нейните романтични новели].</p> <p><i>Име на международна организация на различни езици:</i></p>	

	<p>„Организация на обединените нации“ начин на позоваване на колектив United Nations на български език;</p> <p>'Nations Unies' начин на позоваване на колектив United Nations на френски език;</p> <p>'Nazioni Unite' Начин на позоваване на колектив United Nations на италиански език;</p> <p>'Vereinigste Nationen' начин на позоваване на колектив United Nations на немски език.</p>	
--	<p>LRM-E10 Място (Place)</p> <p><i>Определение:</i> Част от пространството.</p> <p><i>Примери:</i></p> <p>Марс;</p> <p>Италия;</p> <p>Квебек;</p> <p>Клонмакнойс (ирл. <i>Cluain Mhic Nóis</i>) [място в Алтон, Ирландия, където се намира ранноирландско монашеско селище].</p>	
--	<p>LRM-E11 Период от време (Time-span)</p> <p><i>Определение:</i> Определен отрязъкот времето.</p> <p><i>Примери:</i></p> <p>2015-03-01 [един ден от Грегорианския календар във формат ГГГГ-ММ-ДД];</p> <p>XX век;</p> <p>Династията Мин;</p> <p>Бронзова епоха.</p>	

Приложение 2. Добри практики за реализиране на електронни библиотеки

Най-значимите примери в областта на интеграцията на документи и обекти, притежание на различни културни институции, са Световната дигитална библиотека* (World Digital Library, WLD) и Европеана** (Europeana). В хода на работата по тези проекти са достигнати забележителни резултати в сферата на каталогизацията на обектите и създаването на многоезикови портали.

Световната дигитална библиотека стартира официално през 2009 г. с подкрепата на Организацията на обединените нации за образование, наука и култура – ЮНЕСКО. Проектът има за цел да предостави безплатен многоезиков онлайн достъп до дигитализирани обекти на културното наследство на страните членки на организацията патрон (днес 195 страни²⁸⁹). Библиотеката съсредоточава и представя различни по тип електронни ресурси като ръкописи, редки и ценни издания, карти, звукозаписи, филми, фотографии, архитектурни планове и мн. др. При реализацията на проекта се акцентира върху значимостта на обектите, достоверността на информацията, и на различни аспекти на качеството на представянето на обектите и на достъпа до тях²⁹⁰.

Страните на Европейския съюз активно провеждат дейности по интеграцията на информационни ресурси на различни културни институти чрез съвременните информационни технологии. Перспективна международна инициатива, координирана от италианското министерство на културата, обединява усилията в тази насока – проект ATHENA („Достъп до мрежи, свързани с европейското културно наследство”). В него участват със свои дигитални колекции институции от 20 страни членки на Европейския съюз (ЕС), включително и Азербайджан, Израел и Русия. Успешното реализиране на проекта в периода 2008–2011 г. се свързва със създаването на **Европейската дигитална библиотека Европеана**. Тя функционира от 2008 г. (на база опита на проекта за европейска библиотека (The European Library) като веб портал за цифрова културна продукция, който свързва дигитални сбирки на културното наследство на държавите от

* <https://www.wdl.org/en/>

** <http://www.europeana.eu/portal/en>

²⁸⁹ UNESCO. List of the 195 Members [online]. [Viewed 15.07.2017]. Available from: <http://www.unesco.org/new/en/member-states/countries/>

²⁹⁰ Нишева-Павлова, Мария. *Създаване на хетерогенни цифрови библиотеки...* [онлайн]. с. 20.

ЕС²⁹¹. Порталът представя в съвременното информационно пространство разнообразни фрагменти на културно наследство на Стария континент – изобразително изкуство, мода, музика и фотография и др. „Europeana е дигитална библиотека, музей и културен архив”²⁹².

Europeana предоставя възможност ползвателите да търсят в мащабен „централизиран каталог” различни обекти на културно наследство. Създаден е достъп до т.нар. „сурогати” на обекта, които съдържат метаданни, необходими за търсене на езика, на който той е предоставен - неголямо изображение (thumbnail), корица, обложка или заглавна страница на книга или документ, кратък аудио- или видеозапис. Налична е хипервръзка към институцията притежател, където обектът е представен с пълно описание, висококачествено изображение, пълен текст или цялостен аудио- или видеофайл.

Продължаващата инициатива EuropeanaConnect, финансирана от Европейската комисия по програма от областта на електронните библиотеки (eContentplus Programme), разширява функционалните възможности и услуги на Europeana, като осигурява:

- многоезиково търсене и навигиране чрез инструменти за превод;
- семантично обогатяване на цифровото съдържание и създаване на нови връзки между обектите – предпоставка за по-голямата достъпност и възможността за многократно и широко използваемо съдържание;
- интерфейс за мобилни устройства за Europeana;
- интегриране на географски информационни услуги (GIS) и възможност за поръчване на електронни книги по заявка;
- включване на нови услуги, изисквани от потребителите;
- развиване на ключови компоненти от инфраструктурата на Europeana, свързани с инициативата за разработване и популяризиране на стандарти за оперативна съвместимост, за отворен достъп (OAI, Open Archives Initiative) и др.²⁹³

²⁹¹ **Ролята** на Europeana за цифровия достъп, видимостта и използването на европейското културно наследство – *Заключения на Съвета (31 май 2016 г.)* [онлайн]. [Прегледан 07.07.2017]. Достъпно от: <http://data.consilium.europa.eu/doc/document/ST-9643-2016-INIT/bg/pdf>

²⁹² **Христова**, Боряна, **Панчева**, Цветанка, **Чакалов**, Андрей. *Опитът на Националната библиотека „Св. Св. Кирил и Методий“ в дигитализирането на книжния фонд* [online]. [Прегледан 09.07.2017]. Достъпно от: <http://www.bvu-bg.eu/sem/sem8/BHristova.pdf>

²⁹³ **About**: EuropeanaConnect is a Best Practice Network funded by the European Commission within the area of Digital Libraries of the eContentplus Programme [online]. In: *EuropeanaConnect*. [Viewed 24.05.2018]. Available from: <http://www.europeanaconnect.eu/about.php>

До 2025 г. Европейска планира да предостави достъп до напълно дигитализираното културно наследство на Европа. Към настоящия момент тя интегрира близо 53,5 милиона обекти в цифрова форма от различен тип медия – изображения, текст, видео- и аудио и др. Като общ проект Европейска насърчава формалните споразумения между музеи, архиви, аудио-визуални архиви и библиотеки, така че те да си сътрудничат в предоставянето на културно наследство и в постигането на устойчивост на своите колекции. България участва с цифрово съдържание чрез институциите: Народна библиотека „Св. св. Кирил и Методий”, Национален археологически музей, Център по Тракология към Института по Балканистика на БАН, Институт по Литература на БАН, Централна библиотека на БАН, Национална академична библиотечно-информационна система (НАБИС), Регионална библиотека - Пловдив, Регионална библиотека - Варна и др.

Електронни библиотеки на национални библиотеки

Ключово значение за общественото развитие имат институциите на паметта (библиотеки, музеи и архиви), отговорни за събирането, съхранението и разпространението на националното културно и научно наследство, тъй като владеят интелектуалния капитал и технологиите, необходими за обработката и преобразуването на информацията.

„В новите реалности на съществуване и разпространение на информационно-документалните потоци тези институции запазват и дори увеличават своята неизменна роля не само като пазители на националната памет, реализирана в човешкото знание и достижения, но и в други аспекти, свързани със стойността и автентичността на електронната информация”²⁹⁴.

Със своите уникални фондове, с професионалните библиографски и нормативни данни за тях библиотеките и особено националните библиотеки имат много важна роля в съвременното информационно пространство. Националните библиотеки са наточени да предоставят точни библиографски метаданни за културното и интелектуалното производство на държавата²⁹⁵.

²⁹⁴ Стефанова, Радослава Марио. *Електронните ресурси с отдалечен достъп...*, с. 9.

²⁹⁵ Svensson, Lars G. National Libraries and the Semantic Web: Requirements and Applications. In: Prasad, A.R.D, Madalli, Devika P. Eds. *International Conference on Semantic Web and Digital Libraries*[online]. 2007, pp. 101–108. [Viewed 24.05.2018]. Available from: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.113.2756&rep=rep1&type=pdf>

Затова накратко ще представим полезния опит по създаване на електронни ресурси от национални библиотеки.

Националните библиотеки осъществяват функциите по събиране, обработване, организиране, съхраняване и предоставяне за ползване от гражданите на печатни и други произведения, издадени на територията на страната. Те поддържат уникални колекции от ръкописи, архивни документи, редки и ценни старопечатни книги; колекции от специални издания (картографски и графични издания, музикални издания и други); колекции от печатни и други произведения на националната книжнина и чуждестранна литература във всички области на знанието.

Често обаче достъпът до произведения, съхранявани в националните библиотеки, е затруднен именно поради тяхната уникалност. Решение на проблема с ограничения достъп до документални паметници се намира в цифровото им представяне и съхранение. Преимуществовата на електронната форма на информацията са добре известни – надеждност и компактност при съхраняване, възможност за оперативно разпространяване и широко използване. Процесът по дигитализация на документално наследство обикновено започва с най-ценните сбирки на националната библиотека. Тази тенденция се наблюдава във всички държави.

Десетилетие след първите стъпки по създаването на електронни библиотеки (средата на 80-те г. на 20 в.) започва активно разработване на програми за тяхното развитие в САЩ и Европа. През 1993 г. правителството на САЩ публикува програма за създаване на национална информационна инфраструктура, която се разглежда като единна мрежа от комуникационни мрежи, компютри, бази от данни, която ще даде достъп на гражданите до огромно количество информация (National Information Infrastructure: Agenda for Action²⁹⁶). Един от съществените елементи в информационната инфраструктура на страната става проектът за създаване на Национална електронна библиотека (National Digital Library Program, NDLP, 1995²⁹⁷) на Конгресната библиотека.

В Европа по същото време така нареченият „доклад Бангеман” („Европа и глобалното информационно общество”²⁹⁸) набелязва предизвикателствата и стъпките на

²⁹⁶ *The National Information Infrastructure: Agenda For Action Information Infrastructure*[online]. Washington, D.C.: Executive Office of the President, 1993, с. 26, с.1. [Viewed 16.04.2017]. Available from: https://archive.org/stream/04Kahle000911/04Kahle000911_djvu.txt

²⁹⁷ **National** digital library program. In: *The Library of Congress* [online]. [Viewed 16.04.2017]. Available from: <http://memory.loc.gov/ammem/dli2/html/lcndlp.html>

²⁹⁸ **Bangemann** report: Europe and the global information society. In: *European Commission* [online]. CORDIS. Last updated on: 1994-07-07. Record Number: 2730. [Viewed 16.04.2017]. Available from: http://cordis.europa.eu/news/rcn/2730_en.html

Европа към информационното общество. Европейската комисия определя създаването и развитието на електронни библиотеки като общоевропейска политика.

Още през 1995 г. някои европейски страни (напр. Великобритания, Германия, Унгария) предприемат стъпки по електронизация на библиотечно-информационната дейност. Днес националните библиотеки на много страни поддържат свои електронни библиотеки и са намерили решение на ред предизвикателства, свързани с дигитализацията – критерии за подбор на документите за дигитализиране, надеждно съхранение, начин на достъп. Не по-малко важни са и проблемите, произтичащи от авторските права върху дигиталните обекти.

Като опитни в решаването на тези предизвикателства се сочат националните библиотеки на САЩ, Великобритания и Франция.

Библиотеката на Конгреса в САЩ (Library of Congress) в днешно време се приема за най-голямата библиотека в света²⁹⁹. Нейният фонд е универсален – без ограничения по тема, вид или тип на материалите, националност или език. Две трети от материалите, които придобива всяка година, са на езици, различни от английския. Библиотеката към 2016 г. притежава близо 165 млн. библиотечни единици: над 38 млн. печатни издания (книги, периодика, технически отчети и др.), повече от 125 млн. специални материали (ръкописи, аудио, видео, карти, филми, нотирана музика, фотографии и др.) и 3 млн. други ресурси (включително машинночетими)³⁰⁰. Фондът на Конгресна библиотека добавя ежедневно около 10 000 документа (в най-широк смисъл).

Дигиталните колекции на Конгресна библиотека са несравним световен информационен ресурс по обем и обхват. Те се измерват количествено в 7 петабайта информация. Това са 295 тематични колекции³⁰¹, включващи повече от 15 млн. единици (от тях 150 000 са книги). Въпреки впечатляващата цифра това представлява по-малко от 10% от общия размер на фонда, за който, за да се дигитализира напълно дори в един идеален сценарий за дигитализация с темп от 3 до 5 петабайта годишно, ще са необходими десетилетия³⁰².

²⁹⁹ **Fascinating Facts**. Statistics. In: *The Library of Congress* [online]. [Viewed 20.07.2017]. Available from: <https://www.loc.gov/about/fascinating-facts/>

³⁰⁰ **General Information**. In: *The Library of Congress* [online]. [Viewed 20.07.2017]. Available from: <https://www.loc.gov/about/general-information/>

³⁰¹ **Digital Collections**. In: *The Library of Congress* [online]. [Viewed 20.07.2017]. Available from: <https://www.loc.gov/collections/?c=160&sp=2>

³⁰² **Chayka**, Kyle. *The Library of Last Resort*. In: *n+1* [online]. [Viewed 28.06.2018]. Available from: <https://nplusonemag.com/online-only/online-only/the-library-of-last-resort/>

Британската национална библиотека (The British Library) притежава над 150 милиона документи на почти всички езици – ръкописи, карти, вестници, списания, отпечатъци и рисунки, ноти, патенти, звукозаписи, марки и др. Всяка година нейният фонд се обогатява с около 3 мил. екземпляри, голямата част - получавани по депозит³⁰³.

Електронните колекции на Британската национална библиотека са представени чрез порталът British Library Collection guides:

- „*Manuscripts*” – колекцията съдържа ръкописи и архивни документи от фонда на Британската библиотека. Сред най-ценните и интересните са: „Harley Golden Gospels” (ръкопис от 9 век, написан със златно мастило), ръкопис на „Беоулф” (Beowulf – англосаксонска епическа поема, датирана между VIII и XI век, първите преписи съществуват от X — XI в.), „Leonardo da Vinci’s Notebook”, „the Petit Livre d’Amour and the Golf Book”, „Диамантената сутра”, „Четвероевангелието на цар Иван Александър” и „Видинското евангелие”.

- „*British Newspaper Archive*” – архивът съдържа 20,5 млн. дигитализирани страници от вестници, излизали от началото на XVIII в. насам. Създаването на архива е резултат от партньорско споразумение между Британската библиотека и онлайн организацията „findmypast” за дигитализиране на 40 милиона страници за период от десет години.

- „*Electronic Resources*” – точка за достъп и търсене в около 800 електронни ресурси като бази от данни, колекции от електронни книги и електронни списания с абонамент за Британската библиотека. Повечето от тях са с лицензионни ограничения и са достъпни само за регистрирани потребители в читалните на библиотеката.

- „*Social welfare*” – свободен достъп до подбрана висококачествена колекция от пълни текстове на изследвания, проучвания и оценки, парламентарни доклади, правителствени документи и др., свързани със социалната политика на страната.

- „*Sounds*” – колекция от 90 хиляди звукозаписа (музика, драма и литература, устни разкази, природни звуци и др.), подбрани от огромния музикалният фонд (3,5 млн. записи) на библиотеката.

- „*Theses*” – съдържа 450 000 дисертации, защитени в над 120 образователни институции във Великобритания. Около 1/3 от тях са достъпни пълнотекстово. Нивото

³⁰³ **Digital** Collections. In: *British Library Labs: Experiment with our digital collections* [online]. [Viewed 28.06.2018]. Available from: <http://labs.bl.uk/Digital+Collections>

на достъп се определя от отделните институции*. Всеки месец се добавят около 2000 записа с пълни текстове на дисертации.

- „*Turning the Pages*” – колекция от 30 000 електронни книги, които могат да се четат и слушат.

- „*Treasures in full*” – ресурсът представя шест специални колекции с висококачествени дигитални версии на изключително редки и безценни документи в световен мащаб:

- ✓ „**Shakespeare In Quarto**”: 107 екземпляра от различни издания на 21 пиеси на Шекспир;
- ✓ „**Caxton's Chaucer**” - двете издания на „Кентърбърийски разкази” на Дж. Чосър, отпечатани през 1476 и 1483 г. от първия английски печатар Уилям Какстън;
- ✓ „**Gutenberg Bible**” - двата екземпляра на Библията (хартия и велум), отпечатана от Й. Гутенберг през 1455 г.;
- ✓ „**Magna Carta**” (или „Великата харта на свободите”) – документ на латински език от 1215 г., подписан от английския крал Джон Безземни, с който се гарантират и защитават правата и интересите на феодалната аристокрация, бароните и гражданите от действията на представителите на кралската власт.
- ✓ „**Renaissance Festival Books**” – представя 253 материала, части от книги, които описват различни празници: бракосъчетания, погребения, коронации, кралски посещения и други важни събития и фестивали от ренесансова Европа. Най-ранната книга е от 1499 г.
- ✓ „**Malory's Arthurian Manuscript**” – представя страници от ръкописа на сър Томас Малори за крал Артур, известен като „Уинчестърски ръкопис” („*Winchester Manuscript*”), по-късно отпечатан от Уилям Какстън със заглавие „*Le Morte Darthur*”.

Националната библиотека на Франция (*Bibliothèque nationale de France, BnF*) води началото си от кралската библиотека, основана от крал Шарл V през 1368 г. Тя е разширена от крал Луи XIV и е отворена за публиката през 1692 г. Франция е първата страна, която въвежда закон за депозита (1537 г.). Днес сбирката на френската национална библиотека се състои от 15 милиона документи – печатни издания, ръкописи, отпечатъци, фотографии, карти и планове, монети, медали, звуко- и видеозаписи и др. В

* <http://ethos.bl.uk/Home.do>

духа на енциклопедизма са представени всички области на интелектуалната дейност, изкуството и науката. Годишно фондът се допълва с приблизително 150 000 документа чрез депозитни екземпляри или чрез покупка, дарения и обмен.

Електронната библиотека на „BnF”, „Gallica”^{*} представя и осигурява достъп до над 4,2 млн. дигитални обекта: 1 869 805 периодични издания, 521 319 книги, 93 829 ръкописи, 47 430 нотирана музика, 128 971 карти и др. Сред най-ценните дигитализирани „съкровища” на библиотеката са: „Кумранските ръкописи” (или „Ръкописите от Мъртво море”), които представляват фрагменти от 981 различни текста, датирани между III век пр. Хр. и I век; „Камеята на Тиберий” (или „Голямата камея на Франция” (фр. *Grand Camée de France*) изработена от петслоен сардоникс през 23 г.; Папирус от I–II в., наричан „Alexander Papyrus” (или *Romance Papyrus*), който представлява илюстриран фрагмент от неидентифицирано гръцко произведение; Богато украсена богослужебна книга, наречена „Сакраментарий от Дрого” (фр. *Sacramentaire de Drogon*), създадена 845–855 г., която се смята за значим образец на изкуството на украсата на книги и мн. др.

Един от най-мощните проекти за създаване на електронна национална библиотека (Национална електронна библиотека, НЭБ), която да предоставя интернет достъп до дигиталните документи, съхранявани в библиотеки, музеи и архиви се провежда в Русия³⁰⁴. Проектът е по инициатива на Руската държавна библиотека (Российская государственная библиотека) и се реализира от Министерството на културата на Русия. Първият етап на проекта започва през 2004 г., а през 2013 г. е предоставен потребителски достъп до **Националната електронна библиотека на Русия**³⁰⁵. В нейния каталог има 39 634 901 библиографски записа на общо 4 361 992 електронни документа, значителната част с изтекли авторски права^{**}.

Национални библиотеки на по-малки европейски държави, използващи малко разпространен език, също са постигнали значителни резултати в областта на създаването и предоставянето на свободен достъп на уникални електронни колекции. Като примери можем да посочим Финландия и България.

* <http://gallica.bnf.fr/>

³⁰⁴ НЭБ е разгледана подробно от проф. Е. Павловска, затова в изследването представяме само най-съществени данни виж: Павловска, Елена. От библиотеката на Ярослав Мъдри до националната електронна библиотека на Русия. В: *През призмата на информационната наука*. София: Унив. изд. „Св. Климент Охридски”, 2016, с. 329–339.

³⁰⁵ **О проекте**. В: *НЭБ: Национальная электронная библиотека* [онлайн]. [Прегледан 28.05.2018 г.]. Достъпно от: <https://нэб.рф/about/>

^{**} *Информацията е към 28.06.2018 г. Виж: НЭБ в цифрах. Пак там.*

Националната библиотека на Финландия* (фински (suomi) Kansal liskirjasto, шведски (svenska) National biblioteket) е юридически депозиториум и получава копия от всички печатни, аудио и визуални материали (с изключение на филми), произведени или разпространявани в страната. Освен това тя е натоварена да събира и съхранява информация с художествена и научна стойност, публикувана в интернет.

Финландската национална библиотека е дигитализирала 11 512 168 страници към 2017 г., представени в следните колекции: „Вестници”, „Стисания”, „Ефимерни издания”, „Книги”, „Карти и атласи”, „Карти на Норденишелд”, „Класическа литература”, Дисертации, защитени в Кралския университет в Турку (Turun akatemian väitöskirjat) в периода 1642 - 1828 и др. Сред най-интересните колекции се сочат:

- „*Raita – musiikkia vanhoilta äänilevyiltä*” – колекция от 482 музикални произведения от периода 1901–1953 с изтекли авторски и сродни права. Националната библиотека е дигитализирала над 10 000 музикални произведения, които могат да бъдат слушани в читалните на библиотеката.

- „*Fragmenta membranae-kokoelma*” – колекция от 9 319 дигитализирани листа пергамент от около 1500 различни средновековни ръкописа, свързани с историята на най-стария град във Финландия - Турку (от XIII век).

- „*Fenno-Ugrica*” – колекция от дигитализирани над 1500 монографии и теченията на повече от 100 заглавия на периодични издания на различни езици от Уралското езиково семейство.

Националната библиотека на България започва процеса по дигитализация на съхраняваното от нея книжовно наследство сравнително по-късно - през 2006 г. Въпреки това резултатите днес са впечатляващи – до началото на 2016 г. са сканирани и представени на свободен достъп чрез електронната ѝ библиотека около 330 000 файла.

Електронната библиотека е структурирана в зависимост от вида и хронологичния обхват на документите в 27 колекции като: „Славянски ръкописи”, „Българският XIV век” „Славянски кирилски печатни книги XV–XX”, „Български старопечатни книги (1806–1878 г.)”, „Гръцки ръкописи”, „Други чуждоезикови ръкописи”, „Чуждоезикови старопечатни и редки книги XV–XXI”, „Български исторически архив (3 различни колекции- „Документи”, „Кореспонденция”, „Кондики”), „Фотографска колекция – Портрети”, „Печатна графика”, „Картографски издания”, „Продължаващи издания 1844–1878”, „Продължаващи издания 1878–1944” и др.

* <https://www.kansalliskirjasto.fi/en>

Наблюдението, което бе проведено върху електронните информационни ресурси, създадени от националните библиотеки, очерта няколко важни извода:

– По силата на това, че се поддържат от националните библиотеки, тези ресурси са резултат от съвместната дейност на министерства и различни културни и образователни институции (библиотеки, музеи, архиви, университети и др.).

– Реализираните проекти за създаване на електронни библиотеки се характеризират с универсалност по съдържание и вид на документите и материалите и липса на времеви и езикови ограничения.

– Проектите на националните библиотеки за дигитализация на фондовете се подчиняват на идеята за обединяване в глобална информационна система като например *The European Library** - портал, откриващ достъп до ресурсите на 48 национални библиотеки от Европа.

Електронни библиотеки на висши училища

Сред останалите институции, ангажирани със създаване и разпространение на информация и знание – научни институти, информационни центрове, библиотеки, архиви, музеи и др. под. особено място заемат висшите училища. Техните основни информационни ресурси – образователни и научни, в настоящия момент се създават по правило в дигитална форма. Те имат голям опит в работата с електронни каталози, създавани в помощ на автоматизираните библиотечно-информационни системи, които в определена степен могат да се считат за прототип на електронните библиотеки. За висшите училища предоставянето на свободен достъп до собствени информационни ресурси има важни позитивни последици. Най-общо това е форма на популяризиране и обществена експертиза за дейността им.

Създадените електронни библиотеки с научно съдържание в голяма част са тематични и съдържат основно дигитални версии/аналози на печатни издания. Електронните библиотеки на висшите училища, в частност на университетите**, обединяват множество научни направления. Спецификата на дейността на висшите училища – подготовка и повишаване на квалификацията на специалисти и развитие на

* <http://www.theeuropeanlibrary.org/tel4/>. От 31 декември 2016 г. съдържанието на електронната библиотека не се актуализира.

** Университетът обучава по широк кръг специалности от професионални направления в поне три от четирите основни области на науката – хуманитарни, природни, обществени и технически.

науката, културата и иновационната дейност³⁰⁶, прави задължителни следните изисквания към електронните библиотеки:

– Да бъдат пълнотекстова информационна система, която акумулира всички собствени информационни ресурси, а също и външни, до които висшето училище има достъп;

– Тъй като висшите училища разполагат с информационни системи с различно предназначение – за електронно обучение, за административна дейност и др., електронната библиотека може да бъде част от системата им за информатизация. Съхраняваната информация в електронната библиотека може да се използва от друга подсистема на висшето училище или негово подразделение. Например списък на публикации, изобретения и други научно-приложни разработки при процедури за заемане на академични длъжности би могло да бъде кумулиран от електронната библиотека;

– Висшите училища създават разнообразни електронни колекции. Отделните несвързани между себе си електронни колекции не могат да удовлетворяват всички потребности на ползвателите. Необходимо е обединяването им в единна информационна система. Това позволява систематизиране и отчитане на създадените информационни ресурси, разширява спектъра на предоставяните услуги: общо търсене във всички колекции, предоставяне на единна точка за взаимодействие с други системи и др.

Електронната библиотека е ефективен способ за повишаване качеството на информационното обслужване в сферата на образованието. Тя може да реши основни задачи пред всяка университетска библиотека:

- съхранение на фонда;
- книгоосигуреност на учебния процес (един електронен документ може да се използва едновременно от неограничен брой ползватели);
- достъп до информацията, независимо от нейното физическо местонахождение, и по всяко време.

От гледна точка на модерното образование, което предполага контекстно-зависима, адаптивна и персонализирана доставка на образователни услуги и учебно съдържание, изследователи разглеждат електронните библиотеки в ролята им на

³⁰⁶ **Закон** за висшето образование. В: *Lex.bg* [online]. [Прегледан 25.05.2018]. Достъпно от: <https://lex.bg/bg/laws/ldoc/2133647361>.

хранилища на електронно учебно съдържание като компонент на съвременното образователно пространство³⁰⁷.

Историческите функции на институцията Университет – да обучава и развива научна дейност, предопределя някои важни особености на университетските електронни библиотеки. Тъй като те обслужват две основни направления – обучение и научни изследвания, това предполага различни по съдържание и нееднородни по типа информация колекции (текст, цифрови данни, изображения и т.н.). На университетската електронна библиотека трябва да се гледа като на платформа за провеждане на научни изследвания и на комуникация.

За да подкрепим тези общи изводи ще приведем примери из опита на водещи висши училища в създаването и по-нататъшната работа с електронни библиотеки. Важно е да подчертаем, че тяхната практика е свързана с концепцията за Отворен достъп, формирана от началото на века с големи международни инициативи*. Тя насърчава предоставянето на свободен онлайн достъп до научна информация, който е безплатен за крайния потребител и може да се ползва многократно. Единият от способите за реализиране на отворен достъп е създаването на институционални хранилища³⁰⁸.

Институционално хранилище или репозиториум се определя като:

*„дигиталните колекции на научната продукция на университет или друга научна институция, чиято цел е да съхраняват и улесняват отворения достъп до тази продукция. За целта те спазват общ набор от технически стандарти, което означава, че излагат метаданните (библиографските елементи като имена на автори, институционална принадлежност, дата, заглавия на статии, абстракти и т.н.) на базата на общ протокол (OAI-PMH**), което ги прави оперативно съвместими. Така*

³⁰⁷ **Орзова**, Даниела. Извличане на знания от данни в образователно пространство. В: *Годишник на БСУ* [online]. XXXIV, 2016, с. 5–19 [Прегледан 25.05.2018]. Достъпно от: <http://research.bfu.bg:8080/jspui/handle/123456789/650>

* *Декларациите от: Будапеща 2002 г. (Budapest Open Access Initiative), Бетесда 2003 г. (Bethesda Statement on Open Access Publishing) и Берлин 2003 г. (Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities).*

³⁰⁸ **Глинков**, Сергей С., **Клисурова**, Марина, **Радева**, Румяна. *Обзор на ресурсите с отворен достъп в България* [online]. [Прегледан 21.04.2018]. Достъпно от: https://www.researchgate.net/publication/320827255_A_Review_of_the_Current_Situation_of_Open_Access_in_Bulgaria

** *OAI-PMH (OAI Protocol for Metadata Harvesting) – протокол, чрез който хранилищата правят своите метаданни достъпни за услуги и приложения, основани на откритите стандарти HTTP и XML.*

съдържанието на всички хранилища се индексира от уеб търсачки като Google и Google Scholar, създавайки онлайн бази данни с отворен достъп³⁰⁹”.

Предложеното определение не противоречи на концепцията за електронната библиотека, затова при прегледа на добри практики ще бъдат включени и някои институционални хранилища.

Кеймбриджкият университет (University of Cambridge) е един от най-старите университети в света³¹⁰, основан през 1209 г.

Университетската библиотека (*Cambridge University Library*) съществува от шест столетия (създадена 1416 г.) и е натрупала приблизително 15 млн. библиотечни единици, сред тях – карти, ноти, ръкописи и рисунки и др. Библиотеката на Кеймбриджкия университет е една от петте библиотеки, която получава депозитни екземпляри на издания и материали, излезли или разпространявани във Великобритания и Ирландия.

Представява децентрализирана мрежа с 15 филиала и общо 114 специализирани библиотеки към училища, факултети, департаменти или колежи, които притежават поотделно между 30 000 и 200 000 книги. Отделните библиотеки са и хранилища на редки и ценни книги. Например „Wren Library” на Тринити колеж разполага с книги, издадени преди XIX в., „Parker Library” при Corpus Christi College е едно от големите хранилища на средновековни ръкописи (повече от 600 на брой)³¹¹.

Мисията на Университета е да допринесе за развитието на обществото като стимулира образованието и научните изследвания на най-високо международно ниво³¹².

Електронната библиотека на Кеймбриджкия университет е създадена в периода 2010-2014 г. в резултат от финансово дарение от д-р Леонард Полонски (£1,5 мил.) за дигитализация на колекции на библиотеката и предоставянето им на широк достъп чрез „Cambridge Digital Library”*. Тази щедра подкрепа дава възможност на библиотеката да развие техническата инфраструктура и да създаде значително дигитално съдържание, представено в уникални колекции, съдържащи текст, изображения (вкл. подвижни), аудио и др. Стремешът е да се предостави свободно достъпно съдържание за използване

³⁰⁹ Пак там.

³¹⁰ **Early records.** In: *University of Cambridge* [online]. About the university. [Viewed 03.09.2017]. Available from: <https://www.cam.ac.uk/about-the-university/history/early-records>

³¹¹ **Libraries Directory.** In: *University of Cambridge* [online]. [Viewed 26.05.2018]. Available from: http://www.lib.cam.ac.uk/libraries_directory/libraries_directory_n.cgi

³¹² **The University's mission and core values.** In: *University of Cambridge* [online]. [Viewed 26.05.2018]. Available from: <http://www.cam.ac.uk/about-the-university/how-the-university-and-colleges-work/the-universitys-mission-and-core-values>

* <https://cudl.lib.cam.ac.uk/>

в преподаването и при научните изследвания. Електронната библиотека на Университета в Кеймбридж служи като платформа за изследвания в областта на хуманитарни науки в рамките на университета и извън него. Тя притежава 47 колекции, в които са организирани безценни документи с историческо значение, уникалност, красота или завладяващо съдържание. Най-ценните от тях са представени в колекцията „Съкровищницата на библиотеката”. Гордост за библиотеката е сбирката от ръкописи на Исак Нютон.

Електронната библиотека предоставя различни възможности за достъп до документите. Преглеждане по колекции, опростено търсене по дума, която може да бъде от пълния текст или от метаданните, разширено търсене по библиографски данни, например автор, заглавие, предмет, година и др.

Университетската библиотека отговаря за поддържането на институционално хранилище Apollo*, което съхранява резултатите от научните изследвания на академичния състав на университета.

Оксфордският университет (*University of Oxford*) е първият университет в англоговоряния свят. Съществува от 1167 г.³¹³

Библиотеката на Оксфордския университет носи името на един от своите благодетели – сър Томас Бодли (Sir Thomas Bodley, 1545–1613), наричана често „Бодлеанска библиотека” (*Bodleian Library*). Създадена е през 1602 г., а няколко години след това (1610 г.) придобива право да получава депозитни екземпляри³¹⁴. Бодлеанската библиотека притежава близо 12 млн. библиотечни единици и специални колекции от ръкописи, папируси, карти, музика, изобразителни издания, ефимерни издания и др. Това я прави втората по богатство библиотека във Великобритания и най-голямата университетска библиотека в страната. Всеки един от 38-те колежа на Оксфордския университет разполага със своя библиотека, която е част от единната библиотечно-информационна система (*Oxford Libraries Information System*)³¹⁵.

* <https://www.repository.cam.ac.uk/>

³¹³ **Introduction** and history [online]. In: *University of Oxford* [Viewed 03.08.2017]. Available from: <https://www.ox.ac.uk/about/organisation/history?wssl=1>

³¹⁴ **History** of the Bodleian. In: *Bodleian Library & Radcliffe Camera: University of Oxford* [online]. [Viewed 25.05.2018]. Available from: <http://www.bodleian.ox.ac.uk/bodley/about-us/history>

³¹⁵ **Libraries**. In: *University of Oxford* [online]. [Viewed 25.05.2018]. Available from: <https://www.ox.ac.uk/research/libraries?wssl=1>

Колекциите на Бодлеанската библиотека са изключителни, както като обект на научни изследвания, така и като историческо и естетическо богатство, което има стойност за широк кръг потребители. Макар и академична библиотека, 40% от нейните потребители са външни за университета. Целенасочените и дългогодишните усилия да бъдат разкрити част от ценните колекции на библиотеката за обществено ползване – учене, преподаване или изследване, довеждат до създаването на електронна библиотека *Digital Bodleian** през 2015 г. В нея към момента са представени 650 000 свободно достъпни дигитални обекта**, организирани в 10 рубрики: първопечатни книги, ефимерни издания, история и политика, карти, съвременни архиви и ръкописи (XIX–XIX в.), Изток, щампи и портрети, наука и естествена история, западни ръкописи, и представени в 44 колекции.

Основната цел на електронната библиотека е да интегрира отделните дигитални колекции на университета като предложи общ удобен и бърз потребителски интерфейс за търсене и навигиране. Задачите, реализирали целта, са формулирани така:

- стандартизиране на метаданните на отделните колекции;
- преобразуване на всички изображения във един формат (jpeg2000) и преминаване към надеждна мащабируема инфраструктура за съхранение;
- осигуряване на възможност репозиториумът на метаданни да бъде достъпен за авторизирани приложения за събиране на метаданни в съответствие с протокол OAI-PMH;
- осигуряване на възможност потребителите да правят етикети (тагове) и анотации и да обединяват съдържание в собствени виртуални колекции, които могат да споделят с други потребители;
- възможност потребителите да експортират библиографски описания (метаданни) и изображения³¹⁶.

Висшето техническо училище в Цюрих, Швейцария (*Eidgenössische Technische Hochschule Zürich*, ETH Zurich) е създадено през 1855 г.³¹⁷ Понастоящем училището е сред най-влиятелните в областта на инженерните и техническите науки в международен мащаб.

* <http://digital.bodleian.ox.ac.uk/>

** *Информацията е актуална към август 2017 г.*

³¹⁶ **About** the Library. In: *Digital.Bodleian* [online]. [Viewed 25.05.2018]. Available from: <https://digital.bodleian.ox.ac.uk/about.html>

³¹⁷ **Portrait** 2018: „Where the future begins”. In: *ETH Zurich* [online]. [Viewed 25.05.2018]. Available from: <https://www.ethz.ch/en/the-eth-zurich/information-material/information-brochure.html>

Библиотеката на Висшето техническо училище в Цюрих е най-голямата научно-техническа библиотека в Швейцария. Тя оказва силно влияние върху технологичното развитие на всички библиотеки в страната. Тематичния профил на библиотеката е свързан със следните научни области: архитектура, строителство, инженерни науки, природни науки и математика, управление и социални науки.

Електронната библиотека на училището* функционира като уеб портал, който осигурява търсене и достъп до разнообразни електронни ресурси (вестници, списания, карти, геоданни, изображения, аудио-визуални материали, дисертации и др.), представени на различни сайтове или платформи. Тя е част от общото информационно пространство на „ETH Zurich”, като предоставя услуги, свързани с електронно публикуване, съхраняване на цифрови данни и информация и осигуряване на достъп. Задачите, които изпълнява библиотеката, са следните:

- предоставяне на електронна информация на научното съобщество на Висшето техническо училище в Цюрих и на „външни” потребители;
- разработване и предоставяне на иновативни услуги;
- дългосрочно съхранение на научното и културното наследство на Висшето техническо училище в Цюрих в цифрова форма;
- изпълнение на политиката на училището за инициативата „Отворен достъп”;
- управление на цифровите архиви на училището („ETH Zurich University Archives” и „ETH Board Archives”);
- поддържане и по-нататъшно развитие на електронните колекции³¹⁸.

Сред по-интересните електронни колекции са: „Einstein Online” – съдържа кореспонденцията на Алберт Айнщайн**; „Short portraits”, която представя знаменити личности, свързани с училището; „E-Pics Image Archive Online” - 3 млн. дигитализирани фотографии и изображения, свързани с темите: история на науката и техниката в Швейцария и на „ETH Zurich”, снимки на лица, сгради, населени места и др.п. обекти, свързани с Швейцария, фотографски архив на „Swissair” и архив на фотографската агенция „Comet Photo AG”.

* <http://www.library.ethz.ch/en/Resources/Digital-library>

³¹⁸ **Function**, collections, statistics. In: *ETH Zurich: ETH Library* [online]. [Viewed 25.05.2018]. Available from: <http://www.library.ethz.ch/en/About-us/Function-collections-statistics>

** Алберт Айнщайн е бил професор по теоретична физика в „ETH Zurich” в периода 1912–1914 г.

Харвардският университет (Harvard University) се намира в град Кеймбридж, щата Масачузетс, САЩ и е основан през 1636 г. Той е най-ранно създаденото висше училище на Американския континент, с най-старата библиотека (от 1638 г.) в страната. Днес библиотечната система на университета се състои от 79 библиотеки с общ обем на фондовете им – 18,9 млн. единици³¹⁹. По данни на Американската библиотечна асоциация (American Library Association, ALA) библиотеката на Харвардския университет е трета по-големина библиотека и най-голямата академична библиотека в САЩ³²⁰.

Университетската библиотека е дигитализирала над 1 милион обекти, освен това разполага с 5,4 терабайта съдържание, изначално създадено в дигитална форма. Електронната библиотека на Харвардския университет HOLLIS* е изградена като всеобхватен, независим, „вертикален“ уеб портал за търсене и откриване на информация от различни информационни ресурси**. Предоставя на потребителя интегрирано търсене в различните информационни системи, поддържани от университета³²¹.

HOLLIS включва книги, списания, изображения, карти, архиви, ръкописи, ноти, музика, видеоклипове, филми и данни от колекциите на университета, от бази от данни и списания, както и от дигиталното хранилище на Харвард за публикации на отворен достъп (Digital access to scholarship at Harvard, DASH)***.

Харвардската електронна библиотека има мисия да подкрепя обучението и преподаването, като се ангажира със създаването, използването, съхраняването и разпространението на знания.³²²

³¹⁹ **Harvard** Library Annual Report FY 2013. In: *Harvard University* [online]. [Viewed 25.05.2018]. Available from: <http://library.harvard.edu/annual-report-fy-2013>.

³²⁰ **The Nation's** Largest Libraries: A Listing By Volumes Held. In: *ALA American Library Association* [online]. [Viewed 25.05.2018]. Available from: <http://www.ala.org/tools/libfactsheets/alalibraryfactsheet22>.

* <https://hollis.harvard.edu/primoxplore/browse?vid=HVD2>

** *Информационната системата е реализирана през 2014 г. и е базирана на т.нар. Discovery Service (чрез софтуера на Ex Libris Primo), който позволява интегрирано търсене. Discovery Service – единен интерфейс, осигуряващ интегриран достъп до множество информационни ресурси (каталози, колекции на издатели и електронни издания, абонаментни бази данни, архивни колекции), за които библиотеката има права. Discovery Service използва обединено, общо индексирание по тема и метаданни. Резултатите от търсенето обикновено се представят по релевантност. [Източник: http://www.abc-clio.com/ODLIS/odlis_d.aspx]*

³²¹ *Повече за системите за интегрирано търсене виж .Янакиева, Елена. Нови системи за обединен достъп до информация. В: Лозанова-Белчева, Елица, състав. През призмата на информационната наука. София: Унив. изд. „Св. Климент Охридски“, с. 193–205.*

*** <https://dash.harvard.edu/>

³²² **Mission & Objectives.** In: *Harvard Library* [online]. [Viewed 25.05.2018]. Available from: <https://library.harvard.edu/objectives-priorities>.

Принстънският университет (Princeton University) е частен университет в САЩ, създаден първоначално като Колеж на Ню Джърси (1746 г.). По-късно е преименуван в Принстънски университет (1896 г.)³²³.

Библиотечната система на университета е изградена от главна библиотека – Мемориална библиотека „Харви Файърстоун“ („Harvey S. Firestone Memorial“), 9 специализирани библиотеки (по архитектура, инженерни науки, музика, изкуство, Източна Азия, Ръкописи и др.) и три книгохранилища. Общият библиотечният фонд се състои от 10 милиона печатни издания, 5 милиона ръкописи, 2 милиона аудиовизуални материали и дигитални ресурси³²⁴.

Библиотеката притежава документи с културно, историческо и художествено значение като фрагмент от „Книга на мъртвите“ (Egyptian Book of the Dead), глави 67-165 от около 1250 г. пр.н.е., екземпляр от отпечатаната в нощта на 4 юли 1776 г. „Декларация за независимост на САЩ“ (The Declaration of Independence), копие от т. нар. Гутенбергова Библия (Biblia Latina, 1456 г.)³²⁵, които са достъпни за потребители от целия свят чрез електронната библиотека на университета.

Електронната библиотека на Принстънския университет* (Princeton University Digital Library, PUDL) дава достъп до над 50 колекции от различни документи: постери, карикатури, книги, ефимерни издания, пощенски картички, папируси, ръкописи и др. Материалите са подбрани да представят уникалния фонд на университетската библиотека и да обслужват изследователските и образователни потребности на общността на Принстънския университет. Като пример сред тях е ръкопис - нотна тетрадка от 112 страници, на Лудвиг ван Бетховен. Дигиталните копия са с висока разделителна способност и качество.

Достъпът до ресурсите в електронната библиотека на Принстънския университет е чрез навигиране (преглеждане) по колекции и чрез търсене. Възможно е опростено търсене (в едно поле по дума/и), контекстно търсене – по елементи от библиографското описание (метаданните): автор, заглавие, предметна рубрика, думи от резюмето на документа, и търсене чрез булеви оператори.

³²³ **Our** History. In: *Princeton University* [online]. [Viewed 25.05.2018]. Available from: <https://www.princeton.edu/meet-princeton/history>

³²⁴ **Library**. In: *Princeton University* [online]. [Viewed 25.05.2018]. Available from: <https://www.princeton.edu/academics/library>

³²⁵ **Encounter** Global History. In: *Princeton University* [online]. Library. [Viewed 25.05.2018]. Available from: <https://www.princeton.edu/academics/library>

* <http://pudl.princeton.edu/>

Масачузетският технологичен институт (*Massachusetts Institute of Technology*, известен с абривиатурата *MIT*) е американски частен университет в град Кеймбридж, щата Масачузетс. Открит е през 1861 г., като обучението започва през 1865 г. Институтът е едно от най-елитните висши училища в областта на техниката и технологиите.

Библиотеката на института, която тогава притежава едва няколко тома от дарение, отваря врати през 1862 г. – три години преди университетът да посрещне първите студенти³²⁶. Понастоящем библиотечната система е изградена от 5 библиотеки по тематичните направления – инженерство („Barker Library”), икономика („Dewey Library”), хуманитаристика („Hayden Library”), музика („Lewis Library”) и изкуство и архитектура („Rotch Library”), няколко по-тясно специализирани библиотеки и архивът на института. Общият им фонд се състои от около 3 млн. тома печатни материали, 17 000 продължаващи издания, 478 онлайн бази данни, над 55 000 електронни списания с лицензионен достъп и над 2,8 млн. специфични документи – микроформи, карти, изображения, нотни издания, звукозаписи и видеозаписи³²⁷.

Качеството, количеството и достъпността на колекциите на библиотеката са от важно значение за академичния успех и научноизследователската продуктивност. Електронната библиотека на Института отговаря на търсенето на електронна информация с колекции като е-книги и е-списания, дигитализирани архиви на продължаващи издания, учебни електронни ресурси, ресурси на отворен достъп, дигитализирани дипломни работи, защитени в института и др. п. Сред специалните електронни колекции са: „Harold Isaacs collection”, „Noam Chomsky collection”, „Herb Romeroy collection”, „Kenneth Hale tapes”.

Библиотеката подпомага и подобрява преподаването, ученето, научните изследвания и иновациите в рамките на Масачузетския технологичен университет и в световен мащаб. Библиотечната мрежа предоставя чрез експертен персонал и съвременни технологии достъп до качествена научна информация. Библиотеката има за цел да развива знанията в рамките на института и извън него като се основава на:

– Разширяване на достъпа до информация - Масачузетският технологичен институт е сред лидерите в инициативата „отворен достъп” за електронните публикации,

³²⁶ MIT Highlights Timeline [online]. In: *MITstory* [Viewed 25.05.2018]. Available from: <http://mitstory.mit.edu/mit-highlights-timeline/#event-first-wind-tunnel>

³²⁷ Massachusetts Institute of Technology Libraries [online]. In: *Wikipedia* [Viewed 03.08.2017]. Available from: https://en.wikipedia.org/wiki/Massachusetts_Institute_of_Technology_Libraries

достъпни в мрежовото пространство, освободени от лицензионни и правни ограничения и безплатни за потребителя. От 2009 г. насам 46% от научните публикации на Масачузетският технологичен институт са споделени открито чрез библиотеката. Те са изтеглени над 8,5 милиона пъти от почти всички точки на света.

– Съхраняване на знанието - Масачузетският технологичен институт е начело на развитието на технологии, не само за достъп и разпространение на информация, но и новаторски технологии за запазването на цифровата информация.

– Осигуряване на социално пространство - Петте библиотеки предоставят възможност за проучвания и съвместна работа като някои от тях са отворени 24 в денонощието.

– Участие в научноизследователска и развойна дейност - Библиотеките полагат непрекъснати усилия, за да се справят със сложни предизвикателства при разработването на нови устойчиви модели за производство, използване, управление, споделяне, запазване и откриване на научна информация в дигиталната ера³²⁸.

Официално декларираната политика е част от визията на института за създаване на достъпно и динамично научно пространство.

Електронната библиотека функционира като портал с интегрирано търсене в ресурсите на института. Сред тях специално внимание заслужава DSpace@MIT*. Репозиториум на отворен достъп с непрекъснато нарастващ обем, който осигурява на научната съобщност стабилно и дългосрочно съхранение на научни изследвания и учебни материали, както и предоставяне им на максимално широк кръг от потребители. Репозиториумът към момента съдържа над 90 000 материала: доклади от конференции, научни статии и студии, препринти, технически доклади, работни документи, набори от данни за научни изследвания. DSpace@MIT се счита за едно от най-големите хранилища на научна информация в света, от което се осъществяват над милион изтегляния на месец³²⁹.

³²⁸ MIT Libraries. In: *MIT Massachusetts Institute of Technology* [online]. [Viewed 25.05.2018]. Available from: <http://web.mit.edu/facts/libraries.html>.

* <https://dspace.mit.edu/>

³²⁹ About DSpace@MIT [online]. In: *MIT Libraries* [online]. [Viewed 25.05.2018]. Available from: <https://libguides.mit.edu/dspace>

Репозиториумът е базиран на DSpace – специално създаден от Института* софтуерен пакет за обслужване на институционални хранилища на отворен достъп или на електронна библиотека.

Университет „Рупрехт-Карлс“ или Хайделбергски университет (Ruprecht-Karls-Universität Heidelberg) е най-старият университет в Германия. Действа от 23 октомври 1386 г.^{330, 331}.

Библиотечната система на Университет „Рупрехт-Карлс“ е изградена от една централна библиотека с най-голям фонд (приблизително 3,2 млн. книги, 10 732 научни периодични издания и около 500 000 други медии като микроформи, аудио, видео и визуални материали) и 83 библиотеки към факултети и катедри, които общо притежават над 3,5 млн. издания³³². Към фонда на библиотеките принадлежат и 6 600 ръкописа, 1 800 инкунабули, 110 500 автографи и колекция от стари карти, картини и фотографии**.

*Електронната библиотека*** (Digitale Bibliothek Universität Heidelberg)* предоставя интегриран достъп до информационни ресурси на университета: е-списания; е-книги и дисертации; бази от данни и справочници; цифрово хранилище на научните публикации на отворен достъп на Хайделбергския университет; системата за електронно обучение; хранилището за изследователски данни на Хайделбергския университет; база от данни от изображения и мултимедия на университета; дигитални колекции от редки и ценни документи.

Азиатски университети. Авангарден подход за управление и интегриране на електронни колекции прилагат група университети в Хонг Конг: The Chinese University of Hong Kong, City University of Hong Kong, The Education University of Hong Kong, Hong Kong Baptist University, The Hong Kong Polytechnic University, The Hong Kong University

* DSpace е създаден в съавторство с Hewlett-Packard Lab (научно изследователска група към HP Inc., Palo Alto, California) през 2002. От 4 ноември 2002 г. се разпространява съгласно условията на лиценза за отворен код на BSD (Berkeley Standard Distribution License).

³³⁰ **History** of Heidelberg University. In: *Universität Heidelberg Zukunft Seit 1386* [online]. [Viewed 26.05.2018]. Available from: <https://www.uni-heidelberg.de/university/history/history.html>

³³¹ **Heidelberg** University Nobel Laureates. In: *Universität Heidelberg Zukunft Seit 1386* [online]. [Viewed 26.05.2018]. Available from: <https://www.uni-heidelberg.de/university/history/nobellaureates.html>

³³² **Geschichte** der Universitätsbibliothek Heidelberg. In: *Universitätsbibliothek* [online]. [Viewed 26.05.2018]. Available from: <http://www.ub.uni-heidelberg.de/allg/profil/geschichte.html>

** Сред тях голяма ценност представлява Манеският кодекс (*Codex Manesse*), срещан също като „Манески песенник“ („*Manessische Liederhandschrift*“), създаден около 1300–1350 г. в Цюрих. Той съдържа 138 миниатюри, изобразяващи средновековни придворни поети, смятани за шедьоври на горнорейнската готическа илюстрация.

*** <http://www.ub.uni-heidelberg.de/helios/Welcom.html>

of Science and Technology, Lingnan University, The University of Hong Kong. Те създават Консултативен комитет на обединените университетски библиотеки (The Joint University Librarians Advisory Committee, JULAC) за въвеждане на обща интегрирана библиотечна система (integrated library system, ILS) за управление на колекции (печатни и електронни, със стратегически акцент върху електронните ресурси и електронните колекции). Общата система съдържа над 18 милиона библиографски записа и от 2017 г. обслужва над 380 000 потребители³³³.

Целта е осемте библиотеки членове на комитета да използват обща (споделена) система за управление на ресурсите си и общи инструменти за търсене чрез облачна услуга. Като ефект се очаква голяма гъвкавост на системата, както и намаляване на разходите за софтуер и хардуер и икономии в дългосрочен план от ИТ-поддръжка за всяка библиотека³³⁴.

Общата библиотечна система на хонконгските университетски библиотеки е система за интегрирано търсене във всички ресурси на библиотеките, основаващо се на централен индекс. Работа на тези служби се базира на следния принцип:

- създаване на единен (обединен) индекс, формиран от индексирането на всички ресурси;
- търсене в този индекс от една точка;
- в резултат от търсенето се достига до хипервръзки към пълните текстове на материалите³³⁵.

Службите за интегрирано търсене (Discovery Services или Web scale Discovery Services) най-общо се определят като взаимодействие между два основни компонента – единен индекс (Central Index) и общ потребителски интерфейс (Discovery Layer). Централният индекс е списък от събрани и обработени (индексирани) метаданни и пълни текстове. А общият потребителски интерфейс служи за интерактивно взаимодействие със системата и с индекса, за откриване и показване на резултати. Той осигурява еднократно търсене в централния индекс, кратко време за отговор, ранжиране по релевантност, възможност за лимитиране или разширяване на резултатите, връзка с пълните текстове³³⁶.

³³³ JULAC Shared ILS. In: *JULAC: Joint University Librarians Advisory Committee* [online]. [Viewed 26.05.2018]. Available from: http://www.julac.org/?page_id=5499

³³⁴ Ibid.

³³⁵ Янакиева, Елена. Нови системи за обединен достъп до информация. В: Лозанова-Белчева, Елица състав. *През призмата на информационната наука*. София: Унив. изд. „Св. Климент Охридски”, с. 195.

³³⁶ Пак там, с. 196–197.

Приложение 3. Предимства и ограничения на информационни езици, отразяващи съдържанието на документи.

Информационен език		Предимства		Ограничения
Класификационни езици	Йерархични класификационни езици	<p>Значителен опит и традиции в използването на йерархичните класификации като средство за организиране на информационните масиви и за провеждане на информационно търсене.</p> <p>Нагледност, видимост на систематиката.</p> <p>Възможност да се управлява пълнотата на получените резултати.</p> <p>Структурираност и мнемоничност на индексите.</p> <p>Независимост от естествен език.</p>		<p>Лишена от гъвкавост структура, обусловена от строгото фиксиране на класификационните деления и реда на тяхната последователност.</p> <p>Йерархичните класификации неизбежно изостават от бързия темп на развитие на науката и терминологията, свързана с тях.</p> <p>Актуализирането на лексиката (вносянето на допълнения и промени) е сложен и отнемащ време процес.</p> <p>Трудности при многоаспектното разкриване на съдържанието на документите и запитванията и сложност при провеждането на многоаспектното търсене на информация.</p> <p>Процесът на индексирание на документите и запитванията се автоматизира много трудно.</p> <p>Интелектуалното индексирание може да бъде субективно.</p>
	Нейерархични класификационни езици	Предметни класификации	<p>Простотата на подреждане (азбучен ред).</p> <p>Разширяемост – лесно се въвеждат нови термини, които отразяват развитието на науката и техниката.</p>	<p>Отразяването на тесни понятия налага използването на дълги и сложни предметни рубрики.</p> <p>Отразяването на парадигматичните отношения между лексикалните единици изисква сложна система от препратки, което усложнява структурата на езика.</p>
		Фасетни класификации	<p>Възможност за многоаспектно представяне на предмета (темата) на документа.</p> <p>Компактност и обзримост на таблиците на фасетната класификация.</p> <p>Възможност за изграждане на групи, според всяка комбинация от признаци и извършване на информационно търсене по всяка комбинация от фасети.</p>	<p>Трудоемкост при разработването.</p> <p>Необходима е висока квалификация и опит на индексаторите.</p>
Словесни езици	Език на ключовите думи		<p>Възможност за търсене по най-нови термини, а също и по тясно тематични, специфични термини, които не са отразени в контролирани речници.</p>	<p>Високи изисквания към потребителя при формулиране на предписанието за търсене. Трябва добре да познава терминологията в областта.</p>

			<p>Удобство на търсенето, особено за специалисти, познаващи терминологията в предметната област.</p> <p>Възможност за търсене по едни информационно запитване в различни информационно-търсещи системи.</p> <p>Напълно автоматизиране на процеса по индексирание.</p> <p>Отсъствие на субективни фактори при индексирането на документите и запитванията.</p>	<p>Необходимо е използването на сложни логически операции при информационното търсене.</p> <p>Висока степен на информационен шум, в резултат на грешна координация на ключови думи.</p> <p>Ниско качество на качество на търсенето по широкотематични запитвания.</p> <p>Ниско качество на търсенето при несъвпадение на лексиката (в документа и запитването).</p>
	Дескрипторен език		<p>Значителна помощ при провеждане на търсенето, особено за неспециалист в предметната област.</p> <p>Възможност за прецизиране на информационното търсене чрез парадигматичните отношения.</p>	<p>Загуба на информация при търсене на лексика невключена в контролирания речник.</p>

Приложение 4. Съпоставка на полетата на MARC 21 и елементите на Dublin Core с определители³³⁷.

MARC 21 полета	DC Елементи	DC определители
541\$c	Метод на придобиване	
310\$a	Честота на издаване	
521	Предназначение	
100, 110, 111, 700, 710, 711\$e	Участник	вид на интелектуалната отговорност \$e.
720\$e		
255, 034	Обхват	Пространствен
522		
650\$z, 651, 662		
751, 752		
043\$c, 044\$c	Обхват	Пространствен ISO3166
651 if ind2=7 & \$2=tgn		Пространствен TGN
033\$a	Обхват	Времеви
533\$b		
260\$c if precedes date	Дата	на Copyright
542\$g		
260\$c\$g	Дата	на създаване
533\$d		
008/07-10	Дата	на публикуване
260\$c		

³³⁷ Източник: MARC to Dublin Core Crosswalk, достъпно от: <https://www.loc.gov/marc/marc2dc.html>

046\$j	Дата	на модифициране	
046\$m\$n	Дата	на валидиране	
520 if ind1=# or 3	Описание	Резюме	
505	Описание	Съдържание	
300\$a	Формат	Размер	
533\$e			
856\$q			ИМТ
340\$a			Медиа
856\$u	Идентификатор	URI	
008/35-37	Език	ISO369-2	
041 with no \$2		ISO639-2	
041 with \$2=iso639-3		ISO639-3	
041 with \$2=rfc1766		RFC1766	
041 with \$2=rfc3066		RFC3066	
041 with \$2=rfc4646		RFC4646	
561		Историческа справка за притежание	
530	Връзка	Има формат	
776\$n\$t			
530\$u	Връзка	Има формат	
776\$o		URI	
774\$n\$t	Връзка	Има част	
774\$o	Връзка	Има част URI	
775\$n\$t	Връзка	Има вресия	
775\$o	Връзка	Има вресия URI	

530	Връзка	Формат (IsFormatOf)
776\$н\$т		
530\$у	Връзка	Формат (IsFormatOf)
776\$o		URI
440, 490, 800, 810, 811, 830	Връзка	Част от
760, 773\$н\$т		
760, 773\$o	Връзка	Част от URI
510	Връзка	Цитиран
785\$н\$т	Връзка	Заменен от
785\$o	Връзка	Заменен от URI
775,786\$н\$т	Връзка	Версия на
775, 786\$o	Връзка	Версия на URI
780\$н\$т	Връзка	Заменя
780\$o	Връзка	Заменя URI
538	Връзка	Изисква
506\$a\$d	Права	Права за достъп
540\$a\$d		
542\$d	Права	Притежател на права
786\$o	Източник	URI
082	Предмет	Десетична класификация на Дюи (DDC)
050	Предмет	Класификация на Конгресна библиотека (LCC)

600, 610, 611, 630, 650 if ind2=0	Предмет	Предметни рубрики на Конгресна библиотека (LCSH)
600, 610, 611, 630, 650 if ind2=2	Предмет	Медицински предметни рубрики (MeSH)
060	Предмет	Класификация на Националната медицинска библиотека САЩ (NML)
655 if ind2=7 & \$2=tgn	Предмет	Тезаурус на географските наименования (TGN)
080	Предмет	Универсална десетична класификация (UDC)
245	Заглавие	
130, 210, 240, 242, 246, 730, 740	Заглавие	Алтернативно заглавие
Leader06, Leader07	Тип	DCMI Type Vocabulary
655 if \$2=dct		