

СОФИЙСКИ УНИВЕРСИТЕТ „СВ. КЛИМЕНТ ОХРИДСКИ“
ФИЛОСОФСКИ ФАКУЛТЕТ
КАТЕДРА „БИБЛИОТЕКОЗНАНИЕ, НАУЧНА ИНФОРМАЦИЯ И КУЛТУРНА
ПОЛИТИКА“

Мария Асенова Аврамова

ЛИНГВИСТИЧНА ОСИГУРЕНОСТ НА ЕЛЕКТРОННИТЕ БИБЛИОТЕКИ

АВТОРЕФЕРАТ

на

ДИСЕРТАЦИОНЕН ТРУД

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

Дисертационният труд е обсъден и допуснат до публична защита с решение на Катедра „Библиотекознание, научна информация и културна политика” към Философския факултет на Софийския университет „Св. Климент Охридски”.

Научно жури

1. Проф. д-р Оля Харизанова – Софийски университет „Св. Климент Охридски”
2. Проф. д-р Елена Павловска – Университетът по библиотекознание и информационни технологии
3. Доц. д-р Милена Миланова – Софийски университет „Св. Климент Охридски”
4. Доц. д-р Крадимира Иванова – Институт по математика и информатика, БАН
5. Доц. д-р Александра Дипчикова – Национална библиотека „Св. св. Кирил и Методий”

Резервни членове

1. Доц. д-р Елена Янакиева – Национална библиотека „Св. св. Кирил и Методий”
2. Доц. д-р Елеца Лозанова-Белчева – Софийски университет „Св. Климент Охридски”

Дисертационният труд е с общ обем от 274 страници и е структуриран в увод (7 с.), основен текст в три глави (200 с.), заключение (3 с.), списък на използваната литература (24 с.) и приложения (34 с.). Основният текст съдържа 13 таблици, 21 фигури и 8 илюстрации. Списъкът на използваните литературни източници се състои от 304 заглавия, от които 88 на кирилица, 216 на латиница.

Публичната защита на дисертационния труд ще се състои на 20.09.2018 г. от 14 ч. в Катедра „Библиотекознание, научна информация и културна политика”, Философския факултет на Софийския университет „Св. Климент Охридски”, бул. „Цариградско шосе” 125, блок 1, ет. 4, ауд. 45 на заседание на Научно жури, назначено със заповед № РД 38-465/06.07.2018 на Ректора на СУ „Св. Климент Охридски”.

Материалите по защитата са в Катедра „Библиотекознание, научна информация и културна политика”, Философския факултет на Софийския университет „Св. Климент Охридски”.

СЪДЪРЖАНИЕ

I. ОБЩА ХАРАКТЕРИСТИКА НА ДИСЕРТАЦИОННИЯ ТРУД	4
Актуалност и значимост на проблема	4
Обект, предмет, цел, задачи и хипотези на изследването	5
Изследователски подход и методи на изследване.....	6
II. СТРУКТУРА И СЪДЪРЖАНИЕ НА ДИСЕРТАЦИОННИЯ ТРУД	8
Структура на дисертационния труд.....	8
Съдържание на дисертационния труд	8
IV. СПРАВКА ЗА ПРИНОСИТЕ В ДИСЕРТАЦИОННИЯ ТРУД	31
V. ПУБЛИКАЦИИ, СВЪРЗАНИ С ДИСЕРТАЦИОННИЯ ТРУД	32

I. ОБЩА ХАРАКТЕРИСТИКА НА ДИСЕРТАЦИОННИЯ ТРУД

Актуалност и значимост на проблема

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 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 Бюрос, 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.

Обект, предмет, цел, задачи и хипотези на изследването

Научната 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 изследване 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.

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 библиотека.
- 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.

Формулирана е следната **изследователска 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.

Изследователски подходи и методи на изследване

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.

II. СТРУКТУРА И СЪДЪРЖАНИЕ НА ДИСЕРТАЦИОННИЯ ТРУД

Структура на дисертационния труд

Дисертационният труд е в обем от 274 страници, от които 245 научен текст и включва въведение, три глави и приложения. Основният текст съдържа 13 таблици, 21 фигури и 8 илюстрации. За съставянето на дисертация са използвани 304 литературни източници, от които 88 на кирилица и 216 на латиница. В преобладаващата си част източниците са в електронна форма и уебресурси – 249, 55 са традиционни носители на информация, от тях 11 са книги, а 45 са статии, студии, доклади или други съставни части.

Дисертацията е придружена от четири приложения: Приложение 1. Йерархия на обектите в IFLA LRM: определения и примери; Приложение 2. Добри практики за реализиране на електронни библиотеки; Приложение 3. Предимства и ограничения на информационни езици, отразяващи съдържанието на документи; Приложение 4. Съпоставка на полетата на MARC 21 и елементите на Dublin Core с определители.

Съдържание на дисертационния труд

Глава 1. *Електронните библиотеки в съвременното информационно пространство* 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 (т. 1.1). 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”¹. Отсъствието на общоприето определение се свързва с

¹ Харизанова, Оля. *Библиотеките и мрежовото общество: Ефекти и трансформации*. София: Унив.изд. „Св. Климент Охридски“, 2010, с. 230.

неясните физически форми и интензивното развитие на този „феномен“². Коментираното понятие, по мнението на специалистите, е „широко“ и „разнопланово“³. 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 достъп⁴) на библиотеката като институция. 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 (т. 1.1.1.). Вторият подход при разглеждане на същността на електронната библиотека акцентира върху нейната реализация като автоматизирана информационна система. Редица изследователи следват този поход и дават свои дефиниции. Въпреки това няма общоприето определение на понятието „електронната библиотека“.

За целите на изследването бе изведена следната собствена дефиниция: ***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 лице),

² Земсков, А. И., Шрайберг, Я. Л. *Электронная информация и электронные ресурсы*. Москва: Фаир, 2007, с. 332.

³ Пак там.

⁴ Закон за обществените библиотеки [онлайн]. В: *Lex.bg*. [Прегледан на 12.12.2016]. Достъпен от: <http://www.lex.bg/bg/laws/ldoc/2135636021>

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 (т. 1.1.2).

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 потребителите⁵.

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.⁶

Общоприета класификация на електронните библиотеки, която да отчита разнообразните особености и характеристики на различните електронни библиотеки не съществува, затова в дисертационното изследване е предложена следната типология на съществуващите електронни библиотеки по различни признаци (т. 1.2):

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; адресиращи съдържание – документите и метаданните се събират и обединяват в единно информационно съдържание; смесени – използват двата подхода за съставяне и поддържане на информационното съдържание.

б) Според видовете документи, които се включват в електронната библиотека: - електронни библиотеки, които съдържат само един вид документи или един тип информационни обекти; - 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).

⁵ 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

⁶ Павлов, Радослав. Цифрови библиотеки с културно-историческо ...

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 общо предназначение;

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; учебни – предназначени за поддържане и подпомагане на образованието и обучението; 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; - мемориални – създадени с цел да представят лице или събитие; 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; библиотеки, създавани от институциите на паметта – библиотеки, архиви, музеи и др.; 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).

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). Това се постига чрез спазването на утвърдени спецификации, стандарти и модели за представяне на информацията (т. 1.3).

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

Познаването и съобразяването с референтен модел* на електронна библиотека DELOS DLRM гарантира оперативна съвместимост на електронната библиотека (т.1.3.1). Според референтния модел електронната библиотека се разглежда като трислойна „структура”, която се състои от: *Електронна библиотека* (Digital Library, DL) ; *Система на електронната библиотека* (Digital Library System, DLS); *Система за управление на електронна библиотека* (Digital Library Management System, DLMS). Основните изграждащи елементи на електронната библиотека са: *Организация* (Organization); *Архитектура* (Architecture); *Политика* (Policy); *Съдържание* (Content); *Функционалност* (Functionality); *Потребители* (User); *Качество* (Quality). 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.

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 (т. 1.3.2). Концептуалният референтен модел CIDOC Conceptual Reference Model, (CIDOC CRM)⁷ 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 информация за културното наследство. „Информационната интеграция е фокусът на CRM”⁸.

* *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.*

⁷ **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> .

⁸ **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> .

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 системи⁹. 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.

Концептуален модел „Функционални изисквания към библиографските записи” (Functional Requirements for Bibliographic Records, FRBR)¹⁰. Библиографската информация има разнообразни приложения и широк кръг ползватели в съвременното информационно пространство. Този факт провокира експертите от ИФЛА да проведат 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. Библиографският запис се разглежда като съвкупност от данни, свързани с обектите, описани в библиотечните каталози и националните библиографии. 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.

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

⁹ **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/OntologiskaiTeisingasKonceptcinisModeliavimas/pildoma/Guarino98-Formal%20Ontology%20and%20Information%20Systems.pdf>

¹⁰ **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

основни понятия – „обект” (същност), „връзка” и „атрибут” (свойство). Моделът обхваща библиографските данни, разбирани в широк, по-общ смисъл. IFLA LRM преразглежда и дефинира всички потребителски задачи, обекти, атрибути и връзки, свързани с библиографската информация.

Общият европейски формат за изследователска информация CERIF (Common European Research Information Format)¹¹ е модел, който концептуализира научноизследователската сфера. CERIF обхваща всички аспекти на научноизследователската информация – проекти, лица, организации, патенти, продукти, библиометрия, идентификатори и др. 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. CERIF дава схема, която гарантира оперативна съвместимост.

В дисертационното изследване 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 (т.1.4).

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 информация¹².

В информационното пространство съществуват различни инструментални средства за създаване на електронната библиотека. В настоящата разработка са представени резултатите от проведено за целите на дисертацията проучване в Директорията на хранилищата с отворен достъп OpenDOAR и са предложени 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 (т. 5).

¹¹ CERIF. In: *euroCRIS Current Research Information Systems*. [online]. [Viewed 24.05.2018]. Available from: <https://www.eurocris.org/cerif/main-features-cerif>

¹² 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

В таблица 1 са представени 15 от най използваните в света програмни продукти. Според критериите програмите да бъдат с отворен код и със световно разпространение бяха подбрани 3 продукта.

Таблица 1. Използвани програмни инструменти за електронни библиотеки (ЕБ) с отворен достъп OAI-PMH и разпространението им (към 12.01.2018 г.)

Название на софтуер	Брой ЕБ (OAI-PMH)	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

От проведеното изследване бе установено, че най-използваните програмни продукти със свободно използване и с широко световно разпространение са DSpace, EPrints, Greenstone. Dspace и Eprint са предназначени за създаване преди всичко на институционални репозиториуми, а Greenstone е програмен продукт за електронни библиотеки.

Изводи от глава 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.

В глава 2 **Развитие на електронната библиотека** 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 (т. 2.1). 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 (т. 2.1.1). 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 (т. 2.1.2). 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.

Тъй като няма общоприето определение на понятието „системи за организация на знанията”, за целите на дисертационното изследване е прието определението Дагоберт Соергел¹³: *„Системи за организация на знанието обхваща широк кръг от системи, като онтологии, схеми за описание (метаданни), таксономии, класификации, структури на уеб директории, тезауруси, речници, фолксономии 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“.*

Системите за организация на знанието може да се прилагат 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. Те улесняват управлението и търсенето на информация и знания от хора и машини.

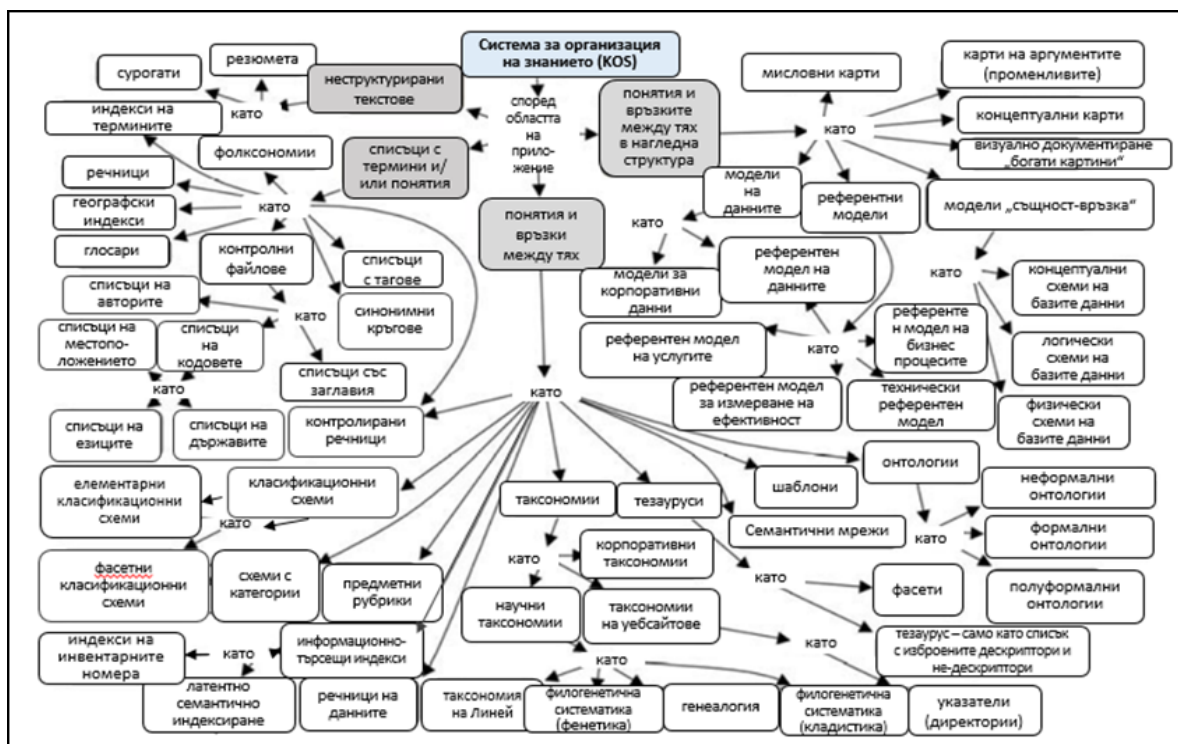
За целите на настоящото изследване е отделено специално внимание върху структурата и типичните свойства на системите за организация на знанията, използвани в библиотеките (т. 2.1.3).

Една от най-изчерпателните типологизации на системите за организация на знанията е предоставена от Ренато Суза и авторски колектив¹⁴ (Фиг. 1).

¹³ **Soergel**, Dagobert. *Knowledge Organization Systems: Overview* [online]. p. 3 [Viewed 20.06.2018]. Available from: <http://www.dsoergel.com/SoergelKOSOverview.pdf>

¹⁴ **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

Фигура 1. Типове системи за организация на знанията ¹⁵



В разработката е предложена типологизация (заимствана от изследванията на Ходж¹⁶), която 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. 2).

Таблица 2. Системи за организация на знанията по Ходж

Категории	Характеристики	Видове
Списъци на термини	Линейни или слабо структурирани 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	Сложно структурирани системи; акцент върху връзките между термините или понятията (еквивалентни, йерархични, асоциативни)	Тезауруси Концептуални карти Тематични карти Семантични мрежи Онтологии

¹⁵ източник: Souza, Renato Rocha, Tudhope, Douglas, Almeida, Mauricio B. Towards a Taxonomy of KOS ...

¹⁶ 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>

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 многократна 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 (т. 2.2).

Разгледана е същността на онтологиите като базово средство за семантично представяне на знанията в уеб (т. 2.2.2.) и е представено тяхното типово разнообразие (т. 2.2.2.1.). 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; логическата сложност и др.

Специално внимание се отделя на приложенията на онтологиите в електронна библиотека (т. 2.2.2.2).

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 (т. 2.2.3).

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. Основната цел на стандарта е да направи възможно публикуването и използване на системите за организация на знанието като свързани данни¹⁷. 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.

¹⁷ SKOS Simple Knowledge Organization System [online]. [Viewed 28.06.2018]. Available from:<https://www.w3.org/2004/02/skos/>

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.

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 и на решенията на социалния уеб за взаимодействие между потребителите. 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. Разгледани са няколко проекта за създаване на семантични електронни библиотеки (т. 2.3).

Изводи от глава 2.

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 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 индексирани. 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, определян 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 съвместно 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 се фокусира 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 (т. 3.1).

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. Лингвистичната осигуреност на електронната библиотека включва различни информационни езици, които имат различно предназначение и роля: 1) цялостно представяне на информационните обекти (документите)* – език на библиографското описание (описателни метаданни); 2)

* 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.

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.

Представена е структурата на информационните езици (т. 3.2.1.): азбука, лексика, отношения, граматика както и технологията на създаването им. Най-важният структурен елемент на информационните системи, от който пряко зависят показателите за ефективност на търсенето (пълнота и точност), са информационните езици¹⁸. 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 (т. 3.2.2.). Проследено е предназначението и спецификите на информационните езици в лингвистичното осигуряване на електронната библиотека (т. 3.3.). Обърнато е внимание върху цялостното отразяване на документите в електронна библиотека се извършва посредством набор от данни, характеризиращи документа по определени признаци, които е прието да бъдат наричани метаданни¹⁹ (т. 3.3.1). Метаданните като централен компонент на лингвистичното осигуряване на всяка информационна система (в частност електронна библиотека) биват няколко вида: описателни метаданни; технически метаданни; метаданни за съхранение; метаданни за права; структурни метаданни. За целите и задачите на настоящото изследване фокусът е поставен върху описателните метаданни, защото именно те служат за представяне на характеристиките на документите в електронна библиотека с цел търсене и идентифициране. Разгледани са и информационните езици за представяне на предмета на документите (т. 3.3.2). За представяне на предмета на документите и за тематичен достъп в библиотечна традиционна или електронна система се използват класификационните езици. Класификационни езици могат да бъдат изградени като йерархични класификации и нейерархични класификации. Детайлно са представени *йерархичните класификации*, които се характеризират с качества като дълбочина и обем и *нейерархичните класификационни езици*, които обхващат езика на предметните рубрики и фасетните класификации.

¹⁸ Гендина, Н. И. Цит. съч., с. 36.

¹⁹ Коголовский, М. Р. Метаданые в компьютерных системах. В: *Програмирование*, 2013, № 4, с. 29. Достъпно онлайн на: <http://www.ipr-ras.ru/articles/kogalov13-03.pdf>

В т. 3.3.3. са проследени информационните езици за многоаспектно представяне на съдържанието на документите – посткоординатните езици, които са от словесен тип, използват за изразяване на лексикалните си единици естествен език. Създаването и използването на посткоординатни езици цели да се избегнат някои съществени ограничения на класификационните езици като сложност при индексирание (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.

Втората част на глава 3 (т.3.4.) 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.

Цели на изследването:

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 е проведено 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 пълно текстови 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/собственост с цел извеждане на зависимости, свързани с опита на традиционните библиотеки в разработването и използването на информационни езици. От подбраните 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 подобрите 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 табличен вид.

Анализът на данните от таблицата очертава няколко важни извода:

1. Първият е свързан с честотата (или активността) на използване на лингвистичните средства:

* включително сдружения на лица, например Българска виртуална библиотека „Словото” се управлява от Сдружение с идеална цел „Словото”.

– 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; индексирането 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.

Заключение 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;

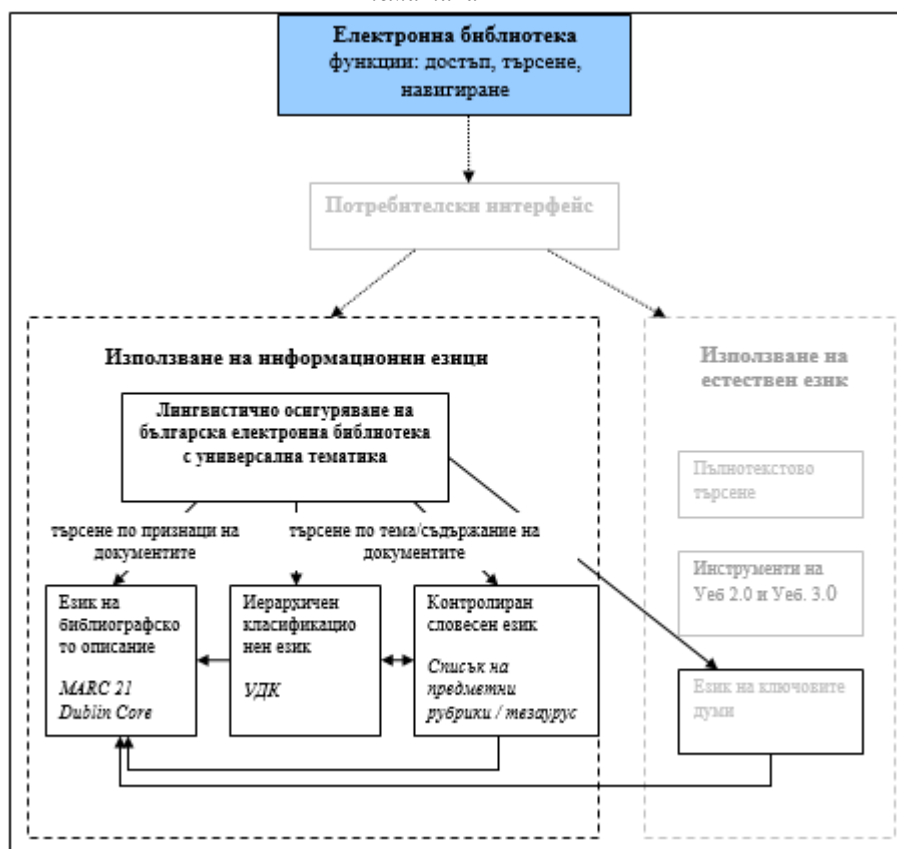
– в малка степен се използват кодовите класификационни езици, които имат предимството да структурират и систематизират информационния масив по удобен начин;

– 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 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.

Резултатите от проведеното изследване са обобщени и са използвани при създаването на модел за лингвистично осигуряване на българска електронна библиотека с универсална тематика (т.3.5.). 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 (фиг.2.).

Фиг. 2. Модел за лингвистично осигуряване на българска електронна библиотека с универсална тематика



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.

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.

Изводи от глава 3.

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 търсеция 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 съдържанието, 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 съкратеният 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.

В заключението са обобщени направените проучвания, анализи, изследвания и изводи по дисертационното изследване.

IV. СПРАВКА ЗА ПРИНОСИТЕ В ДИСЕРТАЦИОННИЯ ТРУД

Научни приноси с теоретичен характер:

1. Чрез анализ на публикации в специализираната научна литература, систематизация и наблюдение е допълнена концепцията на „електронната библиотека” като тип информационна система. Тази информационна система е разгледана относно съхраняване, обработване, структуриране и систематизиране, достъп и търсене на документи и информация и предоставяне на услуги

2. Обобщени и систематизирани са дефинициите на водещи изследователи за „електронна библиотека“ и са представени според избраните от авторката подходи - „библиотековедски“ и електронната библиотека като автоматизирана информационна система.

3. Детайлно е разгледан стандартът Системата за опростена организация на знанията (SKOS) като фактор за развитието на електронните библиотеки и „онтологизирането” на традиционните за библиотеките системи за организация на знанията за използването им като свързани данни в уеб.

Научни приноси с практико-приложен характер:

1. За първи път е направена класификация на информационните езици по различни признаци на български език, подходяща и за обучение на специалисти в областта.

2. За първи път в българската практика и теория се разглеждат възприетите информационни процеси по аналитично-синтетична обработка на документите като лингвистично осигуряване на информационните системи, в частност електронните библиотеки.

3. Изведени са критерии за идентифициране на електронна библиотека, спрямо възприетото в дисертацията определение, които са апробирани в проведено изследване на лингвистичното осигуряване на българските електронни библиотеки в интернет пространството.

4. Разработен и предложен е собствен модел за лингвистично осигуряване на електронна библиотека като специфична информационна система, приложим в практиката на български електронни библиотеки с универсална тематика.

V. ПУБЛИКАЦИИ, СВЪРЗАНИ С ДИСЕРТАЦИОННИЯ ТРУД

1. **Аврамова**, Мария. Концептуалният модел за функционалност на библиографските данни IFLA LRM средство за осигуряване на интеграционни процеси в библиотеки, музеи и архиви. В: *Културната памет във времето и пространството: Сб., посветен на 65-год. юбилей на проф. д.изк.н. Симеон Недков*. София: Университетско издателство "Св. Климент Охридски", 2018, с. 857-870.
2. **Аврамова**, Мария. Електронната библиотека - средство за интегриране на ресурси. Референтен модел на електронната библиотека - DELOS DLRM. В: *Библиотеката - проекции във времето*. Шумен: РБ „Стилян Чилингиров”, 2017, с. 53-66.
3. **Аврамова**, Мария. Езиците за тематично търсене в електронните каталози на българските библиотеки. В: *Годишник на Софийски университет „Св. Климент Охридски”. Философски факултет. Книга Докторанти, Том 2, 2018, с. 260-269 (под печат)*.
4. **Аврамова**, Мария. Лингвистични средства за представяне семантиката на документите и търсене в български електронни библиотеки: изследване и изводи. В: *Сборник материали от международна видеоконференция „Перспективи за културните институции на паметта – библиотеки, музеи, архиви в дигиталната ера*. София-Орел, 10 май 2018 г. (под печат).