



Народна библиотека „П. Р. Славейков“ – Велико Търново



Университет по библиотекознание и
информационни технологии



Великотърновски университет „Св. св. Кирил и Методий“
Катедра „Журналистика и връзки с обществеността“

БИБЛИОТЕКИ ЧЕТЕНЕ КОМУНИКАЦИИ

СБОРНИК С ДОКЛАДИ ОТ
ДВАДЕСЕТ И ТРЕТАТА НАЦИОНАЛНА НАУЧНА
КОНФЕРЕНЦИЯ С МЕЖДУНАРОДНО УЧАСТИЕ

на тема

*Информационното осигуряване
пред предизвикателствата
на изкуствения интелект*

14–15 ноември 2024 г.

Народна библиотека „П. Р. Славейков“
Велико Търново, 2025

*Изданието се осъществява с финансовата подкрепа
на Издателство „Просвета – София” АД*

*Авторите на докладите носят пълна отговорност за
съдържанието, оригиналността, изказаните становища и
цитираните източници и литература.*

- © Доц. д-р Калина Иванова, съставител, 2025
- © Д-р Иван Александров, съставител, 2025
- © Калоян Здравков, графичен дизайн, технически редактор, 2025
- © Авторски колектив, 2025
- © Народна библиотека „П. Р. Славейков“ – Велико Търново, 2025

ISSN: 1313-8138 (print)

ISSN: 2683-0981 (online)

Предговор

„Информационното осигуряване пред предизвикателствата на изкуствения интелект“ бе темата на XXIII национална научна конференция с международно участие „Библиотеки – четене – комуникации“, която се проведе на 14 и 15 ноември 2024 г. в гр. Велико Търново. Традиционно домакин на форума беше Регионална библиотека „Петко Р. Славейков“, а откриването се състоя в Голямата зала на Община Велико Търново, с участието на представители на държавни и общински институции, висши училища, библиотеки, културни организации и фирми от библиотечно-информационния сектор.

Темата на конференцията отразява едно от най-актуалните и динамично развиващи се направления в съвременния свят – навлизането на изкуствения интелект в процесите на създаване, съхраняване, достъп и разпространение на знание. Във време, когато библиотеките и културните институции са изправени пред необходимостта да преосмислят своята роля в обществото на информацията, именно въпросите за информационното осигуряване, критичното мислене, етичните норми и технологичните трансформации се превръщат в централни теми на професионалния и научния дебат.

В рамките на двата конференционни дни бяха представени над двадесет доклада от участници от България, Турция, Сърбия, Португалия, Полша и Германия. Форумът обедини специалисти от различни области – библиотекари, университетски преподаватели, изследователи, експерти по дигитализация и представители на културни институции, които споделиха своя опит и визия за бъдещето на библиотеките в контекста на технологичните промени.

Пленарен доклад изнесе проф. дпн Александра Куманова (УниБИТ, София), която очерта концептуалната рамка на темата чрез представяне на когнитивна платформа за информационно обезпечаване на знаниевото общество – модел, в който библиотеките се разглеждат като авангардни структури в изграждането на интелигентна информационна екосистема.

Сред чуждестранните участници особено впечатление направиха представителите на Университета Бартин (Турция) – доц. д-р Ахмет Алтай и доц. д-р Лале Ождемир Шахин, които разгледаха въпросите за етиката, рисковете и възможностите, произтичащи от навлизането на изкуствения интелект в библиотечното дело. Високо бяха оценени и презентацията на д-р Андреа де Поло Сайбанти (Zeuschel GmbH, Германия), представяща визията на водещата европейска компания за промяната на дигитализационните процеси чрез интегриране на ИИ решения, както и темата на Снежана Ненензич (директор на Народната библиотека в гр. Крушевац, Сърбия) за различните директни приложения на ИИ в дейността на библиотеките.

От българска страна бяха представени редица тематични доклади, които разкриха разнообразието на подходите и практиките в страната. Сред тях се откриха трудовете на проф. д-р Иванка Янкова и нейният екип за ролята на ИИ в трансформацията на приобщаващото образование, на доц. д-р Любомира Парижкова и проф. д-р Таня Тодорова за критичното мислене и предизвикателствата на ИИ, както и на гл. ас. д-р инж. Искра Цветанска за ролята на изкуствения интелект в съхраняването и възстановяването на писмени артефакти. Интерес предизвика и практическото представяне на фирма „Фокус Пойнт“ ЕООД, която демонстрира възможностите на платформата Omeka S и други технологични решения за автоматизация на библиотечните услуги.

Във втората сесия бяха разгледани теми, свързани с комуникацията, образованието и културните политики в условията на дигитална трансформация. Изказванията на д-р Серхио Горжао и Тереза Амарал (Национален дворец Мафра, Португалия), както и на Йоана Чапска (Регионална библиотека „Йероним Лопачински“, Полша), поставиха акцент върху международния обмен на опит и социалното измерение на библиотечните инициативи.

Вторият ден на конференцията бе посветен на иновативни подходи в образованието и културното наследство. Инж. Николай Господинов от РБ „Любен Каравелов“ – Русе, предложи модел на машинно обучение за персонализиране на читателското преживяване. Темите в областта на интелигентните градове, както и исто-

рико-културните изследвания, представени от проф. Димитър Кенанов, доц. д-р Александър Ковачев, Николай Поппетров и Дария Първанова придадоха на форума цялостен и многопластов облик.

Комплимент към участниците и гостите на конференцията беше художественият акцент – концерт на триото Silvan Stâncel от Румъния, който съчета поезията на Матей Вишниец с изтънчена музикална интерпретация, превръщайки научния форум в истински празник на словото, културата и изкуството.

Съорганизатори и дългогодишни партньори на конференцията отново бяха Великотърновският университет „Св. св. Кирил и Методий“ и Университетът по библиотекознание и информационни технологии – София, с подкрепата на Министерството на културата и Община Велико Търново.

XXIII-то издание на националната научна конференция „Библиотеки – четене – комуникации“ потвърди своето място като традиционен форум за научен и професионален обмен, който обединява различни поколения изследователи и практики в общото усилие да се търси баланс между човека, знанието и технологиите.

По традиция докладите от конференцията могат да се ползват и онлайн чрез електронния формат на сборника ISSN: 2683-0981 (online) на сайта на библиотеката в раздел Конференции, достъпен от URL:

<https://libraryvt.com/bg/nauchna-konferentsia-biblioteka-velikotarnovo/2024-naucna-konferencia>

От съставителите

Библиотечните ресурси и услуги в разработки, свързани с изкуствения интелект в рамките на проект КЛаДА-БГ

Иван Крачанов^{1,2} [0000-0002-0430-7953]

Кирил Симов³ [0000-0003-3555-0179]

Димитър Минеv¹

¹Народна библиотека „Иван Вазов“, ул. „Авксентий Велешки“ № 17, 4000 Пловдив

²Институт по математика и информатика, Българска академия на науките,
ул. „Акад. Г. Бончев“, блок 8, 1113 София

³Институт по информационни и комуникационни технологии, Българска
академия на науките, ул. „Акад. Г. Бончев“, блок 2, 1113 София

The role of library resources and services in research related to artificial intelligence within the framework of the CLaDA-BG project

Ivan Kratchanov^{1,2} [0000-0002-0430-7953]

Kiril Simov³ [0000-0003-3555-0179]

Dimitar Minev¹

¹National Library “Ivan Vazov”, 17, Avksentii Veleshki Str., 4000 Plovdiv, Bulgaria

²Institute of Mathematics and Informatics, Bulgarian Academy of Sciences,
8, G. Bonchev Str., 1113 Sofia

³Institute of Information and Communication Technologies, Bulgarian
Academy of Sciences, 2, G. Bonchev Str., 1113 Sofia

Abstract: *The involvement of the National Library “Ivan Vazov” – Plovdiv in the CLaDA-BG project presents a chance for collaboration with research institutions of varied expertise, including those specializing in Artificial Intelligence and Language technologies. The library plays two main roles in this mutually beneficial partnership. By providing digitized texts from its holdings, it assists technology partners in creating text corpora, facilitating computer natural language processing (NLP) and creating generative AI models. The library is also a user of AI applications: improved indexing of text documents, search through queries in Bulgarian, finding similar texts and processing of scanned documents after OCR.*

Keywords: *cultural heritage, semantic web, digital library, digitization, optical character recognition (OCR), generative artificial intelligence.*

Резюме: *Участието на Народна библиотека „Иван Вазов“ – Пловдив в проекта CLaDA-BG дава шанс за сътрудничество с изследователски*

институции с разнообразна експертиза, включително с такива специализирани в областта на Изкуствения интелект и Езиковите технологии. Библиотеката играе две основни роли в това взаимноизгодно партньорство. Предоставяйки дигитализирани текстове от своите фондове, тя подпомага технологичните партньори в създаването на текстови корпуси, улесняване на компютърната обработка на естествен език (NLP) и създаването на модели на генеративния изкуствен интелект, базиран на Големите езикови модели за българския език и за нормативния правопис отпреди 1945. Също така библиотеката е потребител на приложения с ИИ: подобро индексирани на текстови документи, търсене чрез заявки на български език, намиране на подобни текстове и обработка на сканирани документи след OCR.

Ключови думи: културно наследство, семантичен уеб, дигитална библиотека, дигитализация, оптично разпознаване на символи (OCR), генеративен изкуствен интелект.

1, Въведение

Генеративният изкуствен интелект, наричан тук за краткост изкуствен интелект (ИИ), обхваща широк набор от инструменти и методологии базирани на компютърни изчисления, които дават възможност за създаването на съдържание, включително текст, изображения, говор и др. През последните години тези технологии стават все по-общодостъпни и полезни, като това се дължи на усъвършенстването на невронните мрежи, заедно с наличието на достатъчна изчислителна мощност, нужна за обработката на обширните набори от данни в интернет, които се използват за обучение на езикови модели (Leslie & Rossi 2023). ИИ има потенциала да предизвика процеси на глобална трансформация, която има своето отражение и в библиотеките, като предоставя нови възможности за управление на ресурси, за разширяване на обхвата на услугите и адаптиране към развиващите се нужди на потребителите. Библиотеките, като хранилища на знания и центрове за достъп до информация, могат да използват технологии с ИИ за оптимизиране на дейностите си и за подобряване на отношението на потребителите към институцията.

Някои от основните перспективи на инструментите с ИИ са свързани с автоматизирането на трудоемките процеси на катало-

гизиране и генериране на метаданни. Алгоритмите за машинно обучение могат да анализират текстово и мултимедийно съдържание, за да генерират метаданни. Човешката намеса все още е необходима, но с техники за обработка на естествен език (NLP) за автоматична класификация и аотиране на документи, може да се постигне оптимизация на работния процес. Библиотечните колекции стават по-отворени и използвани чрез свързване на ресурсите им в семантични мрежи и графи от знания. По отношение на потребителите, всичко това може да направи откриването на необходимите им ресурси по-успешно, да разполагат с връзки между различни ресурси и да изследват съдържанието чрез концептуални взаимоотношения, вместо чрез обикновени ключови думи. Също така, може да се улесни откриването на модели, тенденции и опорни точки в големи корпуси от текстове, и да се предлага синтезирано съдържание.

Въпреки обещаващия потенциал на ИИ в библиотеките, съществуват редица предизвикателства по отношение на приемането и интегрирането на тези технологии. Притесненията относно загубата на работни места сред библиотечния персонал и необходимостта от адекватно обучение в сферата на ИИ са значителни пречки пред прилагането му. Освен това трябва внимателно да се обмислят етичните последици от използването на ИИ, включително въпроси, свързани с доверието на потребителите в достоверността и изчерпателността на предоставяната от ИИ информация. Потребителите често се чувстват несигурни относно прозрачността и сигурността на системите с ИИ, което може да повлияе на доверието им в библиотечните услуги (Andersdotter 2023). Фактор на несигурност е и догонващото законодателство в сферата. За разрешаването на тези проблеми и успешната интеграция на новите инструменти е необходимо повишаване на квалификацията на библиотечния персонал и ясна комуникация относно спецификите и функционалностите на ИИ (Gupta & Gupta 2023).

2. Библиотеките и семантичният уеб

Увеличаването на обемите дигитално съдържание в онлайн хранилищата на културните институции се превръща в същест-

вено предизвикателство, защото ориентирането в наличната информация и намирането на нужната такава става все по-трудно. Търсенето с ключови думи води до това, че често трябва да бъдат прегледани огромен брой нерелевантни документи и възможност да бъдат пропуснати потенциално ценни източници. В дигиталните колекции на Народна библиотека „Иван Вазов“ (НБИВ) този проблем става все по-актуален, с натрупването на значително количество текст, резултат от дигитализирането на текстови източници, върху които е приложено оптично разпознаване на символи (OCR). Сериозен проблем е точността на това разпознаване – идентифицирането на думите невинаги е коректно, от което следва погрешно включване или пропускане на такива от резултатите от търсенето (Hinze et al. 2016). OCR е много полезен инструмент, но разпознаването много рядко е безупречно и успеваемостта му зависи най-вече състоянието на оригиналния документ, на какъв език е, какъв шрифт е използван и др. На база на опита натрупан в НБИВ, разпознаването на добре запазени оригинали, на съвременен български език, е с процент успеваемост най-често над 99%. По-голямата част от материалите обект на дигитализация обаче, са исторически текстове от периода преди Правописната реформа от 1945 г., което води до намаляване на точността на OCR (Крачанов & Минев 2024). Възможно е ръчно коригиране на грешките след OCR, но това изисква значителни усилия и много време, което прави този метод нецелесъобразен за нуждите на библиотеката и забавя значително темповете за добавяне на ново съдържание.

Автоматизацията чрез ИИ е възможно решение на тези проблеми – както с откриването на нужната информация, така и с коригирането на грешките от OCR. Компютърните системи, за разлика от хората, много по-трудно „разбират“ съдържанието на текста, включително контекстните нюанси. Нуждата от създаване на семантични връзки между понятията и определяне на именуваните същности са основният фокус на задачите за привиждането на езика в такъв, който може да полезен за целите на автоматизацията.

Анотацията е съществен етап в подготовката на текстовете по такъв начин, че да е възможно тяхното обвързване в семантич-

на мрежа. Чрез добавяне на метаданни и допълване на текстовия корпус с информация от по-високо ниво, анотациите позволяват по-ефективно търсене и филтриране на подходящо съдържание. Необходимо е да се установят взаимоотношенията между обектите в текста, за да може да се извлече информация за културното наследство от наличните текстови корпуси, както и тя да бъде категоризирана, организирана и представена по подходящ начин (Stoikov 2021). Възможността за предоставяне на знания в машинночетим формат подобрява оперативната съвместимост между различните приложения на ИИ, което им позволява да комуникират и споделят информация по-ефективно (Zine et al. 2019).

Семантичната мрежа се базира на представянето на знанието чрез онтологии, които са формален набор от понятия в рамките на дадена област и връзките между тези понятия. Онтологията позволяват споделената обработка и извличане на информацията, възможни чрез прилагането на унифицирани идентификатори на ресурси (URI), система за описание на мрежовите ресурси (RDF) и протокол за прехвърляне на хипертекст (HTTP), които съществуват в множество източници на данни. Особено важна в контекста на големите данни, семантичната мрежа дава възможност за по-разширени възможности чрез сложни заявки да се извличат резултати представени по различни начини (Mithun & Bakar 2020).

3. Проект КЛаДА-БГ

НБИВ е част от проекта КЛаДА-БГ, интегриран в рамките на европейските инфраструктури CLARIN и DARIAH, като библиотеката е партньор доставящ съдържание и експертиза. Мисията на проекта е да създаде национална технологична инфраструктура от ресурси за езиковото, културно-историческото наследство и технологии, които да осигурят публичен достъп до тези ресурси, инструменти за обработка на българския език и инструменти за достъп и управление на масивите от данни на културно-историческото наследство за различни обществени задачи, насочени към широката аудитория (КЛаДА-БГ – Мисия 2025). Всички постижения следва да са в полза и да се предоставят на крайните потребители.

За изпълнението на своите задачи КЛаДА-БГ се базира на сътрудничеството на партньори от технологичната сфера с институции и академични учреждения от областта на културата, като е предвидено създаването на мрежа от езикови ресурси и инструменти за културно-историческото наследство, както и разработване на уеб приложения, които да демонстрират потенциала на инфраструктурата по отношение на използваните технологии. Този подход на сътрудничество е от съществено значение за преодоляване на ограниченията пред отделните културни институции, особено по отношение на експертния опит в областта на информационните технологии. Чрез участието си в национална електронна инфраструктура като КЛаДА-БГ, библиотеките могат да получат достъп до по-широк спектър от ресурси и технологии, което може значително да подобри усилията им в полето на дигитализация (Luchev et al. 2021).

КЛаДА-БГ подкрепя развитието на специализирани дигитални инициативи, които отговарят на специфичните нужди на българското културно наследство. Така например, проектът доведе до създаването на средата за съхранение, извличане и съхранение на данни CultIs, която намира реализация за нуждите на редица институции, включително НБИВ (Goynov et al. 2024). Тази инфраструктура не само подобрява съхранението и извличането на цифрово съдържание, но и гарантира, че материалите от културното наследство се съхраняват сигурно и ефективно, и се предоставят на потребителите по лесен за ползване начин, съобразен със съвременните добри практики.

Както стана известно в предната част на доклада, езиковите ресурси и технологиите за тяхната обработка, като извличане и свързване на понятия/ключови думи, именувани обекти и събития, са в основата на развитието на ИИ. Примери за езикови ресурси и технологии, които КЛаДА-БГ предоставя, са следните:

- **BulTreeBank WordNet (BTB-WN):** Български лексикален ресурс, представящ значенията в българския език като група от синоними, между които има мрежа от отношения, отразяващи езиковото (и фактологично) знание.

- WebCLaRK: Интегрирана система за корпуси и речници, която позволява наблюдение на думи и фрази в контекст и достъп до няколко речника едновременно.
- КЛаДА-БГ допринася за концептуалната рамка на цифровите библиотеки чрез интегриране на езиковите и познавателните ресурси в една кохерентна система. Създаването на ориентирана към България графа на знанието в рамките на инфраструктурата КЛаДА-БГ илюстрира тази интеграция, осигурявайки структуриран начин за управление и достъп до езиковото знание (Simov 2019).
- Корпус със описание на събития. Той свързва наименовани същности (хора, организации, локации и други) с конкретни събития, описани в текста. Този корпус е база за модели за автоматичен анализ на текстове с концептуална информация. Резултатът ще се използва за извличане на ново знание към Мрежата от думи и за индексирание на документи (Osenova et al. 2022).

Текущи цели за развитие на проекта в областта на ИИ са в областта на обработката на естествени езици и обучаване на големи езикови модели (large language models – LLM), които да изпълняват различни задачи в КЛаДА-БГ, изграждане на ефективен езиков конвейер (“pipeline”) за български език и генериране на псевдокорпуси за по-нататъшно обучение на LLM. Приложения на тези цели ще позволят автоматично индексирание на документи с именуванни същности и термини от различни онтологични домейни, извличане на знания от текст, генериране на текст от лексикални вериги от ВТВ-WN и др. Примери за работата по проекта в тази насока е разработването на bgGlue – инструмент за оценка на степента на разбиране на българския език от големи езикови модели (Hardalov et al. 2023), както и автоматичното идентифициране на частите на речта и именуваните същности върху съвременен корпус от парламентарна реч ParlaMint 4.0 (Osenova & Simov 2024). Тренирани са базираните на трансформатори езикови модели BERT-WEB-BG и GPTWEB-BG (Marinova et al. 2023).

Приложенията на разработките на КЛаДА-БГ в сферата на изкуствения интелект за подобряване на услугите, предоставени от библиотеките са разнообразни и обхващат всички етапи от подготовката на наличните дигитални ресурси, за да могат да бъдат автоматично обработвани, до конкретни приложения и услуги.

Подобряването на точността на OCR за ресурси на български език е фундаментална необходимост. Проектът КЛаДА-БГ използва мултидисциплинарния опит на партньорските организации за разработване на методологии за OCR, като по този начин улеснява по-доброто боравене и използване на дигитализираните текстове. Резултат в тази насока беше създаването на лексикон на състарени словоформи на базата на правописните правила преди правописната реформа от 1945 г. Резултатите от проведените тестове показват подобряване на разпознаемостта (Kratchanov et al. 2020). В тази насока се включва и нормализацията на извлечения текст, като е трениран модел за превод от стария правопис (отпреди 1945) с точност над 95%.

След първоначалния етап – корекцията на разпознатия текст, следва необходимостта от автоматизираното създаване на мрежа от семантични връзки в дигитализираното съдържание, включване му в графи от знания и свързани отворени данни – обогатяване и търсене на семантично ниво, и обвързване с глобалните масиви от свързани данни.

Едно приложение на ИИ в библиотеките, което се прилага все по-широко, е създаването на чатботове, които дават автоматизирани отговори на въпроси от потребители, като се генерират нова, персонализирана информация. Употребата на ИИ в този контекст е възможност за предоставяне на консултантска услуга на базата на диалози с множество потребители на библиотека едновременно, като същевременно се намалят разходите за обслужване. По този начин се предлага допълнителен канал за комуникация с потребителите във всеки един момент, подобрява се взаимодействието и се показва повишена отдаденост от страна на библиотеките да задоволят информационните потребности на читателите (Калчева 2024).

Предиктивният анализ, подпомогнат от ИИ, може да идентифицира тенденции в търсенето на материали, помагайки на библиотеките да вземат информирани решения относно това какво да предлагат на своите читатели и върху кои ресурси да се съсредоточат при изграждането на дигитални колекции.

Машинният превод също може да намери важно приложение в дейността на библиотеките чрез подобряване на достъпността, разширяване на употребата на предлаганите ресурси и осигуряване на услуга за тези потребители, които не разбират езика. Библиотеките могат да използват ИИ за превод на дигитализирани ресурси, като по този начин да се улесни трансграничната изследователската работа, да се подпомогне изучаването на езици и създаването на метаданни и др. Пример за приложение на тази технология е в Дигиталната библиотека на НБИВ, където в административната част на системата съществува възможност за машинен превод на английски език на въвежданите метаданни за добавяните дигитални ресурси (Goynov et al. 2024) Въпреки че преводите невинаги са перфектни и имат нужда от проверка и понякога корекции, тази помощна функция спестява значително време и усилия от страна на администраторите на системата.

Обобщаването на текстове, силно застъпено направление в сферата на обработката на естествен език, цели синтезиране на значението на определен корпус в по-малък обем. Експоненциално увеличение на обема на данните, достъпни в интернет, усложнява способността на потребителите да се запознаят с всичко, които за тях може да представлява интерес. Чрез автоматизираното обобщаване могат по-бързо да намерят подходяща информация, като първо прегледат резюмето на документа и след това да решат дали да прочетат цялото съдържание (Oliveira & Lins 2024). Автоматизираното определяне на тематиката и близост на текстовете също са възможни приложения на обработката на естествен език, които позволяват препоръка на дигитализирани ресурси на базата на тяхната релевантност на потребителското търсене. Определянето на тематиката може да помогне за създаване на метаданни и извличане на ключови думи.

Автоматизираният анализ на отношението и мненията за продукти, организации, политики и други, е друго полезно и перспективно направление в обработка на естествен език, който се използва за установяване на настроенятията и емоционалния тон, изразени в текст. Различни заинтересовани страни имат нужда да разберат чувствата, позициите или нагласите на отделните лица относно конкретни обекти или събития. Политическите партии се интересуват от проучването на общественото мнение, тъй като представя за предпочитанията на избирателите и наблюдават нагласите за гласуване. Търговските компании се интересуват от обратната връзка от своите клиенти, за да предвидят тенденциите, да формулират насоките и да оценят ефикасността на своите стратегии (Singh et al. 2024). Приложението на автоматизирания анализ на отношението и мненията може да разкрие по нов и полезен начин дигитализираните от библиотеките ресурси. Такъв тип анализ е полезен също за изследователите в областта на хуманитарните науки, за да разберат културните и социалните тенденции, улеснява изследването на обществените настроения и емоционалните реакции към исторически събития, социални движения и културни проявления, като може да се проследи развитието на нагласите във времето. В изследванията на медийната среда, автоматизираният анализ на настроенятията улеснява проследяването на пристрастия, предразсъдъци и обществени реакции на важни събития, на изразените емоции в новинарски статии и реклами, и разкрива както индивидуални, така и обществени тенденции.

Синтезът на естествена човешка реч от текст е друго съществено и бързо развиващо се направление на обработката на естествен език, което може да намери значително приложение в дейността на библиотеките. Особено полезен е за хората с увредено зрение, които се нуждаят от помощ при използване на електронни устройства, за четене на имейли или съобщения. Именно с такава цел от ИИКТ-БАН е разработен синтезаторът за българска реч NeuralSpeechLab (Geneva et al. 2023). Такъв тип инструменти могат да се приложат и върху дигитализираните от библиотеките ресурси, така че те да станат по-достъпни. Синтезът на реч може

успешно да се комбинира с други от гореспоменатите технологии, като например с автоматизирания превод (Sanapala et al. 2024).

4. Заключение

ИИ има потенциала да внесе коренни промени в начин по който библиотеките функционират и предоставят своите услуги. В процеса на интеграция на технологии с ИИ, библиотеките трябва също да се справят с предизвикателства свързани с правото и етиката, доверието на читателите, поверителността на данните и необходимостта от дигитална грамотност сред персонала и потребителите. Обмисленото внедряване на ИИ може да гарантира, че библиотеките ще останат значими институции в дигиталната епоха, насърчавайки разпространението на знания и ученето през целия живот.

Проектът КЛаДА-БГ значително подобрява услугите по дигитализация, предлагани от библиотеките в България чрез подобрени OCR технологии, развитие на специализирани дигитални инфраструктури и интегриране на езиковите ресурси. Тези постижения адресират предизвикателствата, пред които са изправени библиотеките при управлението на дигиталното съдържание, което в крайна сметка води до по-стабилна и достъпна дигитална библиотечна екосистема.

От своя страна, библиотеките са ценни партньори, защото те притежават големи количества текст, и особено такъв върху който няма авторскоправна тежест – важно предимство, защото предоставените данни могат да се използват от хората и машините без ограничения.

Благодарност

Изследването, представено в този доклад/работа, е извършено в рамките на CLaDA-BG, българската национална изследователска инфраструктура за ресурси и технологии за езиково, културно и историческо наследство, интегрирана в рамките на CLARIN EU и DARIAH EU, финансирана от Министерството на образованието и науката на Република България (подкрепя за българската Национална пътна карта за научноизследователска инфраструктура).

Използвана литература:

1. BulTreeBank WordNet [online]. [viewed 14.01.2025]. Available from: <https://clada-bg.eu/en/centers-and-services/language-technologies/btb-wordnet.html>
2. WebCLaRK [online]. [viewed 14.01.2025]. Available from: <https://concordance.webclark.org/>
3. КЛаДА-БГ – Мисия [online]. [viewed 24.01.2025]. Available from: <https://clada-bg.eu/bg/mission.html>
4. ANDERSDOTTER, K. Artificial Intelligence Literacy in Libraries. *Journal of Information Literacy*, 2023, vol. 2, no. 17, pp. doi: 10.11645/17.2.14
5. DANNÉLLS, D., VIRK, S. M. A Supervised Machine Learning Approach for Post-OCR Error Detection for Historical Text. *Linköping Electronic Conference Proceedings*, 2021, doi: 10.3384/ecp184170.
6. GENEVA, D., SHOPOV, G., GAROV, K., TODOROVA, M., GERDЖIKOV, S., МИHOV, S. (2023) Accentor: An Explicit Lexical Stress Model for TTS Systems. *Proc. Interspeech 2023*, pp. 4848-4852, doi: 10.21437/Interspeech.2023-433
7. GOYNOV, M., LUCHEV, D., PANEVA-MARINOVA, D., SENKA, G., et al. CultIS : Web-based Platform for Intelligent Cultural Content Management. *Digital Presentation and Preservation of Cultural and Scientific Heritage*, 2024, vol. 14, no., pp. 19-36, doi: 10.55630/dipp.2024.14.1
8. GUPTA, V., GUPTA, C. Synchronizing Innovation: Unveiling the Synergy of Need-based and Curiosity-based Experimentation in AI Technology Adoption for Libraries. *Library Hi Tech News*, 2023, vol. 40, no. 9, pp. 15-17, doi: 10.1108/lhtn-07-2023-0127.
9. HARDALOV, M., ATANASOVA, P., MIHAYLOV, T., ANGELOVA, G., et al. bgGLUE: A Bulgarian General Language Understanding Evaluation Benchmark. *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 8733-8759, doi: 10.18653/v1/2023.acl-long.487
10. HINZE, A., BAINBRIDGE, D., CUNNINGHAM, S. J., DOWNIE, J. S. 2016. Low-cost Semantic Enhancement to Digital Library Metadata and Indexing: Simple Yet Effective Strategies. *Proceedings of the 16th ACM/IEEE-CS on Joint Conference on Digital Libraries*, Newark, New

- Jersey, USA, 2016 Association for Computing Machinery, pp. 93–102, doi: 10.1145/2910896.2910910
11. KRATCHANOV, I., LASKOVA, L., SIMOV, K. Towards Improving OCR Accuracy with Bulgarian Language Resources. *DHNB Publications*, 2020, vol. 3, no. 3, pp. 115–123, doi: 10.5617/dhnbpub.11226
 12. LESLIE, D., ROSSI, F. ACM TechBrief: Generative Artificial Intelligence [online]. Association for Computing Machinery, New York, NY, United States, 2023. [viewed 22.01.2025]. Available from: <https://dl.acm.org/doi/book/10.1145/3626110>
 13. LUCHEV, D., GOYNOV, M., PANEVA–MARINOVA, D., STOYKOV, J., et al. Synergy of National Cultural Heritage and Technology. *Digital Presentation and Preservation of Cultural and Scientific Heritage*, 2021, vol. 11, pp. 281–286, doi: 10.55630/dipp.2021.11.26
 14. MARINOVA, I., SIMOV, K., OSENOVA, P. Transformer-Based Language Models for Bulgarian. *Proceedings of Recent Advances in Natural Language Processing*, Varna, Sep 4–6, 2023, NCOMA Ltd., Shoumen, Bulgaria, 2023, ISBN:978-954-452-092-2, ISSN:2603-2813.2023, 712-720, doi:10.26615/978-954-452-092-2_077
 15. MITHUN, A. M., BAKAR, Z. A. Empowering Information Retrieval in Semantic Web. *International Journal of Computer Network and Information Security*, 2020, vol. 12, no. 2, pp. 41–48, doi: 10.5815/ijenis.2020.02.05
 16. OLIVEIRA, H., LINS, R. D. 2024. Assessing Abstractive and Extractive Methods for Automatic News Summarization. [online]. *Proceedings of the ACM Symposium on Document Engineering 2024*, San Jose, CA, USA, 2024, Article 12, doi: 10.1145/3685650.3685664
 17. OSENOVA, P., SIMOV, K. Bulgarian ParlaMint 4.0 Corpus as a Testset for Part-of-speech Tagging and Named Entity Recognition. *Proceedings of the IV Workshop on Creating, Analysing, and Increasing Accessibility of Parliamentary Corpora (ParlaCLARIN) @ LREC-COLING 2024*, 2024, p. 30–35.
 18. OSENOVA, P., SIMOV, K., MARINOVA, I., and BERBATOVA, M. 2022. The Bulgarian Event Corpus: Overview and Initial NER Experiments. [online]. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference, European Language Resources Association*, Marseille, France, pp. 3491–3499. European Language Resources Association. [viewed 22.01.2025]. <https://aclanthology.org/2022.lrec-1.374.pdf>

19. SANAPALA, D., CHOUDHARY, K., SHETTY, S. 2024. Multi-Person Speech Curator for Minutes of Meetings along with Meeting Summarization and Language Translation. [online]. *Proceedings of the 2023 7th International Conference on Advances in Artificial Intelligence*, Istanbul, Turkiye, 2024 Association for Computing Machinery, 6–12, doi: 10.1145/3633598.3633600
20. SIMOV, K. Integrated Language and Knowledge Resources for a Bulgarian-Centric Knowledge Graph. *Digital Presentation and Preservation of Cultural and Scientific Heritage*, 2019, vol. 9, pp. 65–74, doi: 10.55630/dipp.2019.9.5
21. SINGH, U., ABHISHEK, K., AZAD, H. K. A Survey of Cutting-edge Multimodal Sentiment Analysis. *ACM Comput. Surv.*, 2024, vol. 56, no. 9, pp. Article 227, doi: 10.1145/3652149
22. STOIKOV, J. Using Conditional Probability for Discovering Semantic Relationships between Named Entities in Cultural Heritage Data. *Digital Presentation and Preservation of Cultural and Scientific Heritage*, 2021, no. 11, pp. doi: 10.55630/dipp.2021.11.7
23. ZINE, O., DEROUICH, A., TALBI, A. IMS Compliant Ontological Learner Model for Adaptive E-Learning Environments. *International Journal of Emerging Technologies in Learning (Ijet)*, 2019, vol. 14, no. 16, pp. 97, doi: 10.3991/ijet.v14i16.10682
24. КАЛЧЕВА, Р. Модел на прилагане на изкуствен интелект (ИИ) в обществените библиотеки в България. *Библиотека*, 2024, бр. 2, стр. 94–99.
25. КРАЧАНОВ, И., МИНЕВ, Д. 2024. Технологични предизвикателства в дигитализацията на културно-историческото наследство в Народна библиотека „Иван Вазов“ в двугодишна перспектива. [online]. *Двадесет и втора национална научна конференция с международно участие „Библиотеки – четене – комуникации“*, на тема „Стандартизация и методология при дигиталната конверсия на културно-историческото наследство“, 16–17 ноември 2023 г. 2024, Народна библиотека „П. Р. Славейков“, Велико Търново. [viewed 13.01.2025]. Available from: <https://libraryvt.com/images/conferences/2023/Proceedings/Ivan-Krachanov.pdf>

Съдържание

Предговор	5
<i>От съставителите</i>	
Критично мислене и информационно осигуряване: предизвикателствата на изкуствения интелект	9
<i>Доц. д-р Любомира Парижкова, проф. дн Таня Тодорова</i>	
<i>Critical thinking and information assurance: the challenges of artificial intelligence</i>	
<i>Assoc. Prof. Dr Lubomira Parijkova, Prof. DSc Tania Todorova</i>	
Информационното осигуряване на знаниевото общество – когнитивна платформа на системно обезпечаване с универсални справочно-информационни ресурси (конвергентен аванпост на библиотечно базирания изкуствен интелект)	21
<i>Проф. д.п.н. Александра Куманова</i>	
<i>Information procurement of the knowledge society – a cognitive platform of a system providing with universal reference-information resources (a Convergent outpost of the library based artificial intelligence)</i>	
<i>Prof. D.Sc. Alexandra Kumanova</i>	
Ролята на изкуствения интелект и технологиите в трансформацията на приобщаващото образование	34
<i>Проф. дфн Иванка Янкова, гл. ас. д-р Камелия Нушева</i>	
<i>гл. ас. д-р Деница Димитрова</i>	
<i>The role of artificial intelligence and technologies in the transformation of inclusive education</i>	
<i>Prof. DSc Ivanka Yankova, Chief. Assist. Prof. Dr. Kamelia Nusheva, Chief. Assist. Prof. Dr. Denitsa Dimitrova</i>	
„Светлата“ и „тъмната“ страна на изкуствения интелект	44
<i>Доц. д-р Любомира Парижкова, Марина Илиева, Светослав Славов, Юрий Тодоров, Марио Тренчев</i>	
<i>The “light” and “dark” side of artificial intelligence</i>	
<i>Assoc. Prof. Dr Lubomira Parijkova, Marina Ilieva, Svetoslav Slavov, Yuri Todorov, Mario Trenchev</i>	

**Ролята на изкуственият интелект в съхраняване
и възстановяване на писмени артефакти..... 58**

Гл. ас. д-р инж. Искра Цветанска

*The role of artificial intelligence in the preservation
and recovery of written artifacts*

Chief Ass. Dr. Eng. Iskra Tsvetanska

**Перспективата на Zeuschel за това как изкуственият интелект
ще промени работния ни процес на дигитализация 63**

Кристиан Хоендорф, Андреа де Поло Сайбанти

*Zeuschel prospective on how AI will change
our digitization workflow*

Christian Hohendorf, Andrea de Polo Saibanti

**Последиците от изкуствения интелект върху библиотеките:
възможност или заплаха? 73**

Доц. д-р Ахмет Алтай, доц. д-р Лале Шахин

*The Repercussions of Artificial Intelligence on Libraries:
An Opportunity or a Threat?*

Assoc. Prof. Dr. Ahmet Altay, Assoc. Prof. Dr. Lale Özdemir Şahin

Използване на изкуствен интелект в библиотеките..... 86

Снежана Ненезич

The application of artificial intelligence (AI) in libraries

Snežana Nenezić

**Потенциалът на българските библиотеки
за преодоляване на различни отклонения
при прилагането на ПР и маркетингови технологии 93**

Доц. д-р Евгения Русинова

*The potential of Bulgarian libraries to overcome
various deviations in the application
of PR and marketing technologies*

Assoc. Prof. PhD Evgeniya Rusinova

**От страница до дебат. Въздигането на читателските клубове
в Полша..... 103**

Йоана Чапска

From Page to Debate. The Rise of Book Clubs in Poland

Joanna Chapska

Омега S и други AI решения в библиотеките: Технологичната еволюция на библиотечните услуги	122
<i>Светослав Николов, Хариета Николова</i>	
<i>Omega S and Other AI Solutions in Libraries: The Technological Evolution of Library Services</i>	
<i>Svetoslav Nikolov, Harieta Nikolova</i>	
Изкуствен интелект за интелигентни градове – оптимизиране на градското пространство.....	136
<i>Дария Първанова</i>	
<i>Artificial Intelligence for Smart Cities – Optimizing Urban Space</i>	
<i>Daria Parvanova</i>	
Библиотечните ресурси и услуги в разработки, свързани с изкуствения интелект в рамките на проект КЛаДА-БГ	152
<i>Иван Крачанов, Кирил Симов, Димитър Минев</i>	
<i>The role of library resources and services in research related to artificial intelligence within the framework of the CLaDA-BG project</i>	
<i>Ivan Kratchanov, Kiril Simov, Dimitar Minev</i>	
За книгопечатника Дякон Кореси.....	166
<i>Проф. дфн Димитър Кенанов</i>	
<i>About the writer Deacon Koresi</i>	
<i>Prof. Dr. Sc. Dimitar Kenanov</i>	
Паралели на достъпността (основни характеристики на библиотеките в полския град Люблин при работата им с хората с увреждания).....	180
<i>Николай Колев</i>	
<i>Parallels of accessibility (basic characteristics of libraries in the polish city of Lublin in their work with disabled people)</i>	
<i>Nikolay Kolev</i>	
Поглед към бъдещето на професионалното образование и обучение	202
<i>Гл. ас. д-р Деница Димитрова</i>	
<i>A look at the future of vocational education and training</i>	
<i>Chief. Assist. Prof. Dr. Denitsa Dimitrova</i>	

Информационното осигуряване в библиотеките – нови възможности	217
<i>Доц. д-р Силвия Станчева</i>	
<i>Information provision in libraries – new opportunities</i>	
<i>Assoc. Prof. Dr. Silvia Stancheva</i>	
За мястото на някои културни институции в изследванията по културна история за периода 1878–1944 г.	225
<i>Николай Поппетров</i>	
<i>About the place of some cultural institutions</i>	
<i>in cultural history studies for the period 1878–1944</i>	
<i>Nikolay Poppetrov</i>	
Политическата комуникация в ерата на изкуствения интелект: Библиотеките като посредници	236
<i>Докторант Чавдар Раев</i>	
<i>Political communication in the age of AI: Libraries as mediators</i>	
<i>PhD Student Chavdar Raev</i>	
Никола Начов – Добавки и оправки към „Български книгопис за сто години“ от академик Александър Теодоров–Балан	249
<i>Доц. д-р Александър Ковачев</i>	
<i>Nikola Nachov – Additions and corrections</i>	
<i>to “Bulgarian Book Writing for a Hundred Years”</i>	
<i>by Academician Alexander Teodorov–Balan</i>	
<i>Assoc. Prof. Dr. Alexander Kovachev</i>	

БИБЛИОТЕКИ – ЧЕТЕНЕ – КОМУНИКАЦИИ

Двадесет и трета национална научна конференция с
международно участие

Велико Търново, 14–15 ноември 2024 г.

Българска

Първо издание

Научна

Съставители:

доц. д-р Калина Иванова

д-р Иван Александров

Формат 60x90/16

Печатни коли 16,5

Печат: Печатница „Сира“ Велико Търново