

Екатерина Лобузина,

директор Института информационных технологий
Национальной библиотеки Украины имени В. И. Вернадского,
доктор наук по социальным коммуникациям,
старший научный сотрудник
Украина, г. Киев
e-mail: lobuzina@nbuv.gov.ua
ORCID: 0000-0003-3371-4029

**СЕМАНТИЧЕСКИЕ ВЕБ-ТЕХНОЛОГИИ
И БИБЛИОТЕЧНЫЕ ПОИСКОВЫЕ СИСТЕМЫ**

Статья посвящена использованию семантических веб-технологий в библиотечно-информационной деятельности. Рассмотрены принципиальные изменения, произошедшие в современной веб-среде, их влияние на эволюцию библиотечных поисковых систем. Определено понятие семантической библиотеки (Библиотеки 3.0), которая использует методы доступа к информации, основанные на сложных семантических отношениях. Освещены изменения подходов к организации электронных библиотечных ресурсов и систем поиска информационных источников. Показана эволюция библиотечных информационных систем от онлайн-овых электронных каталогов к системам выявления знаний. Проанализированы основные характеристики, которым должны соответствовать современные библиотечные каталоги с точки зрения интернет-пользователей: поиск всего контента в одном месте, ранжирование вывода документов на экран, простой поиск по ключевым словам, фасеты для фильтрации результатов поиска, современный интуитивный интерфейс, получение полного текста одним нажатием, подсказки с рекомендациями дополнительных источников информации, расширение библиотечной информации за счет данных из внешних источников. Охарактеризованы подходы использования технологий Семантического Веба в библиотечных поисковых системах: федеративный поиск, связанные данные, фасетный поиск. Продемонстрированы способы использования структурированных открытых связанных данных, таких как WikiData, в библиотечных поисковых системах, что значительно улучшает видимость и доступность библиотечных ресурсов в Интернете. Изложенный материал проиллюстрирован опытом внедрения семантических веб-технологий для поиска информации на портале Национальной библиотеки Украины имени В. И. Вернадского.

Ключевые слова: семантические веб-технологии, информационный поиск, библиотечные поисковые системы, электронные каталоги, библиотечные каталоги нового поколения, системы выявления знаний, связанные данные.

Основной целью семантической библиотеки является организация средств представления информации, которые имеют преимущества по сравнению с сервисами обычных электронных библиотек. Пользователи должны иметь возможность использовать взаимосвязанную информацию о ресурсах в процессе просмотра, фильтрации или нахождения подобных информационных объектов; средства уточнения запроса должны адаптировать свои результаты в соответствии с запросами пользователей; методы доступа к информации должны использовать сложные семантические отношения между результатами. Наконец, семантическая *e*-библиотека должна предлагать различные рекомендательные сервисы, например, на основе контекста или аннотаций. Поисковая система должна предоставлять возможность использовать информацию о различных типах носителей, сложных объектах, текущих и пространственно-временных связанных ресурсах. Важно поддерживать поиск на основе алгоритмов сходства информационных объектов [5]. В библиотековедческих исследованиях термин «семантическая (интеллектуальная) библиотека» употребляется в контексте становления понятия «Библиотека 3.0» (Library 3.0) [3, 6, 20, 24]. Среди основных характерных черт Библиотеки 3.0 следует отметить следующие.

Интеллектуальность. Библиотека 3.0 применяет системы искусственного интеллекта для предоставления интеллектуальных услуг пользователям библиотеки. В частности, она воспринимает запросы на естественном языке, использует библиотечные интеллектуальные инструменты, такие как классификаторы и авторитетные файлы, для устранения неоднозначности текстовых сообщений, применяет гибкие и эвристические стратегии поиска. Это дает пользователям возможность получать информацию на основе ее содержания, а не простого совпадения слов. Используются также онтологии для раскрытия запроса пользователя и предоставления ответов на основе содержания понятия, что делает библиотечные сервисы более интерактивными, точными и удобными.

Организованность и упорядоченность. Информация продолжает расти по объему и сложности, что создает определенные препятствия для ее потребления. Библиотека 3.0 предназначена для того, что-

бы превратить неорганизованную информацию глобальной веб-среды в систематизированную и полезную информацию, где каждый элемент данных атрибутирован и связан с другими с целью обеспечения легкого и быстрого доступа. Это предполагает устранение несогласованности и дублирования данных, которое достигается в процессе сотрудничеством пользователей, экспертов и библиотекарей в направлении создания, унификации, фильтрации и обмена достоверной информацией (в частности, открытыми связанными данными).

Интегрированность. Инструменты Библиотеки 3.0 объединяют различные информационные источники и платформы для создания надежной информационной сети, которая работает беспрепятственно для быстрого, точного и систематического поиска и обмена информацией. Объединяются различные информационные каналы, форматы и среды для обеспечения синергии между еще разрозненными информационными ресурсами и системами с целью усовершенствования доступности, поиска и удобства использования достоверной информации. Объединенные поисковые системы позволяют пользователям без специализированных поисковых навыков мгновенно искать широкий спектр ресурсов и преодолевать проблемы, возникающие вследствие различных интерфейсов входа и разнообразных функций поиска. Библиотека 3.0 способствует поиску объединенных данных на базе открытых стандартов и совместимости метаданных, создает для пользователей единую точку доступа к информации.

Реализация преимуществ сервисов Библиотеки 3.0 предусматривает эффективное использование семантических веб-технологий и онтологий; интеграцию электронных каталогов (далее – ЭК) библиотек с электронными полнотекстовыми ресурсами; создание библиотечных порталов, предоставляющих доступ ко всем библиотечным ресурсам в режиме «единого окна»; дальнейшее развитие виртуальных библиотечных справочных служб, доступных с мобильных телефонов и устройств; активное участие библиотекарей в посредничестве между знаниями и пользователями: приведение, нахождение и предоставление доступа к документированным знаниям. Ведущие специалисты в области информационных технологий отмечают, что сегодня созданы все необходимые методы и технологии для работы с семантическими моделями. Еще в 2013 г. аналитики Gartner объявили семантические технологии одним из самых перспективных трендов развития ИТ-отрасли. В 2018 г. констатировалось, что в следующие пять лет будущее технологий будет

характеризоваться умными устройствами, которые повсеместно будут предоставлять цифровые услуги (интеллектуальные цифровые сети – intelligent digital mesh). Основными характеристиками нового цифрового мира будут: разумность (intelligent), цифровизация (digital), сплетенность (mesh) [16, 17].

Семантическая технология – информационная технология, ориентированная на содержание и связи между данными в процессе их обработки и использования. Чтобы семантическая технология действительно работала в системе, необходимо иметь модель знаний какой-то части мира (активная онтология). Понятие семантической технологии подчеркивает тот факт, что она обеспечивает работу со смысловым содержанием информации. Таким образом, переход от традиционных информационных технологий (далее – ИТ) к семантическим технологиям является переходом от работы с данными к работе со знаниями. Разница между этими двумя терминами заключается в способе использования информации: для восприятия и использования данных человеку необходимо выполнять их интерпретацию и выявлять смысл. Представленные же в электронной форме знания могут восприниматься непосредственно, поскольку они уже выражены с помощью того понятийного аппарата, которым пользуется человек. Кроме того, с такими знаниями (онтологиями) могут выполняться и полностью автоматические операции – получение логических выводов. Результатом этого процесса становятся новые знания [1].

Отметим особенности и преимущества семантических технологий, среди которых ИТ-специалисты IBM выделили основные [18]. Также приведем некоторые примеры, которые касаются библиотечной деятельности.

1. Семантические технологии обеспечивают более высокий уровень абстракции, чем существующие ИТ-технологии.

Происходит это благодаря добавлению к наборам данных контекстной информации. Например, в библиографических данных может добавляться информация о магазинах, где можно приобрести книгу, и ее стоимости; местонахождении книги в библиотеках и как туда добраться; о других изданиях произведения и его переводах; об авторе и других его произведениях; отзывах читателей о книге и сколько раз к ней обращались. Эту дополнительную информацию можно использовать для принятия пользователем эффективных решений в процессе поиска необходимого произведения.

2. Семантическая технология обеспечивает гибкий подход к интеграции информационных ресурсов.

Семантические технологии предлагают лучшую интеграцию и совместимость различных информационных ресурсов, чем традиционные ИТ-системы, в них можно легко вносить изменения. Семантический подход предусматривает возможность установки любых новых связей данных без необходимости повторного перепроектирования всей системы. Например, структура авторитетных файлов и лингвистических справочников (библиотечных классификаторов) должна согласовываться с библиографическими БД так, чтобы ее можно было использовать как справочную информацию в любой ситуации во время информационного поиска: получить справку о различных формах имени автора или названия учреждения, вместо классификационного индекса вывести формулировки на естественном языке и т. п.

3. Семантический Веб: компьютер может сделать много полезной работы за нас, но мы должны подготовить ему данные.

Семантический Веб – это совместная информационная платформа, которую возглавляет организация международных стандартов в сфере веб-технологий – World Wide Web Consortium (W3 C). Его цель – предоставить машинам возможность «понимать» и реагировать на сложные человеческие запросы на основе их содержания. Этот тип «понимания» предусматривает, что соответствующие источники информации должны быть семантически структурированы. Для этого предназначена RDF (Resource Description Framework, далее – RDF) – модель данных для представления знаний в Семантическом Вебе. В контексте библиотечного сообщества еще в 2005 г. была реализована работа по переводу схемы FRBR (Functional Requirements for Bibliographic Records, далее – FRBR) в модель данных RDF [12]. Сегодня полномасштабное внедрение модели данных RDF осуществляется в разработке нового стандарта библиографических записей BIBFRAME, инициированного Библиотекой Конгресса США, где вместо маркеров полей введены специальные RDF-идентификаторы [10].

4. Онтологии являются важнейшим компонентом семантических веб-проектов.

Онтологии являются формальными спецификациями объектов и отношений между ними и лучше всего подходят для представления информационных моделей, необходимых для обеспечения интеллектуальных и прогностических решений. Наиболее пригодными для

преобразования в онтологии в библиотечной деятельности оказались библиотечные классификации и системы предметизации, поскольку именно они представляют в структурированных и связанных терминах определенную предметную область. Описание объектов онтологии осуществляется с помощью семантических метаописаний – особого вида описаний, включающих концептуальное (аннотированное) изложение содержания и сути информации об объекте. Основными требованиями к метаописаниям объектов онтологии являются: полнота представления знаний, закрепленных в онтологии; возможность использования метаописаний объектов в других системах; использование общепринятого стандарта, который совместим с другими системами.

Ключевой в развитии и внедрении современных семантических технологий стала концепция Семантического Веба, разработанная в 2001 г. Т. Бернерсом-Ли [8], изобретателем World Wide Web (WWW), URIs, HTTP и HTML, который определил его как расширение (надстройку) существующей Всемирной паутины с целью представления информации в виде, который позволит лучше понимать ее как людям, так и компьютерам. Семантический Веб рассматривается Т. Бернерсом-Ли в контексте развития Всемирной паутины как WWW второго поколения, который ориентирован на автоматизированную интерпретацию и обработку информационных ресурсов. Основой концепции Семантического Веба стала идея о том, что компьютер пока не может понимать естественной человеческой речи, поэтому нужно использовать язык, который был бы понятен компьютеру. То есть, в идеальном варианте, всю информацию в Интернете необходимо размещать на двух языках: человеческом и компьютерном.

Основу системы организации знаний Семантического Веба образуют связанные данные (*Linked Data*). Связанные данные (термин предложен в 2006 г. Т. Бернерсом-Ли) – гипертекстовая система, назначением которой является публикация структурированных данных, связывание метаданных между собой, представление различных аспектов содержания, выявление и создание перекрестных ссылок между соответствующими ресурсами. *Linked Data Project*, основанный в 2011 г., предлагает общеупотребительные иерархии классов, словари имен, а также помогает владельцам массивов данных объединять их базы знаний в одну всемирную связанную систему знаний. Во многих случаях участники проекта объединяют уже имеющиеся большие базы, помогая друг другу установить соответствие между идентификаторами одного и того же объекта в разных базах. Технология связанных данных имеет следующие преимущества:

- 1) агрегирование данных из распределенных онлайн-первоисточников;
- 2) создание новых связей между ними и их визуализация;
- 3) обогащение данных с помощью ссылок на внешние ресурсы [9, 21].

Масштабным проектом открытых связанных данных (*Linked Open Data*) является *WikiData* (викиданные; <https://www.wikidata.org/>). Предназначение WikiData – представление структурированных данных, содержащихся в материалах всемирной энциклопедии Wikipedia [30]. Практически WikiData выполняет функцию авторитетного контроля, традиционную для библиотечных каталогов. Каждая статья WikiData содержит уникальный идентификатор, название и синонимы заголовка, изображения из WikiMedia, ссылки к связанным понятиям и статьям Wikipedia, варианты представления понятия в Wikipedia на разных языках. Данные из централизованного хранилища могут свободно использоваться различными вики-проектами, их использование можно наблюдать в виде вывода информационных панелей Wikipedia во время Google-поиска. Основу для создания списка статей WikiData составляет англоязычная версия Wikipedia. Так, статья о Национальной библиотеке Украины имени В. И. Вернадского (далее – НБУВ) (<https://www.wikidata.org/wiki/Q592734>) имеет идентификатор Q592734, англоязычный заголовок «Vernadsky National Library of Ukraine», можно обнаружить информацию о наличии статьи об НБУВ в Wikipedia на 15 разных языках.

База данных WikiData предоставляет также возможность использовать структурированные данные в собственных веб-проектах, в частности библиотечных. Технология связанных данных была также использована при создании международного библиотечного сервиса VIAF – Виртуального международного авторитетного файла (*Virtual International Authority File*, далее – VIAF). В частности, статьи WikiData содержат идентификаторы и ссылки к различным сервисам авторитетного контроля VIAF, Международного стандартного идентификатора названий (*International Standard Name Identifier*, далее – ISNI), авторитетного файла Библиотеки Конгресса США, предметных рубрик Национальной библиотеки Франции, авторитетного файла Библиотеки Ватикана и т. п. Все это позволяет перейти к поиску соответствующих информационных источников в библиотеках и агрегированных библиотечных ресурсах, таких как WorldCat. Современные библиотеки активно вклю-

чаются в процесс создания структурированных веб-данных, который является частью движения «Открытых данных» (*Open Data*) – глобальной инициативы по предоставлению доступа к публичным данным с целью их дальнейшего свободного использования и распространения. Любое лицо может их свободно копировать, публиковать, распространять, использовать, в частности в коммерческих целях, в сочетании с другой информацией или путем включения в состав собственного продукта с обязательной ссылкой на источник их получения. Условием любого дальнейшего использования открытых данных является обязательная ссылка на источник информации. В 2016 г. Украина взяла на себя обязательства выполнять принципы Хартии открытых данных [15].

Параллельно с развитием идей структурирования данных веб-серевды также развивались новые подходы к библиографическому описанию, которые подчеркнули важность в библиографическом описании дополнительных точек доступа (тема (предмет), произведение, организация, лицо, место, физические характеристики, аннотация). Эти идеи были сформулированы и закреплены в 1998 г. принятием Международной федерацией библиотечных ассоциаций и учреждений (*International Federation of Library Associations and Institutions*, далее – IFLA) «функциональных требований к библиографическим записям» (*Functional Requirements for Bibliographic Records*, FRBR), которые зафиксировали переосмысленные требования к каталогизационной практике на международном уровне [14]. FRBR, основанные на современных семантических моделях, предложили новую структурированную модель «сущность-связь» (*entity-relationship model*, или сокращенно – ER-модель), которая связывает данные, представленные в библиографических записях, с потребностями пользователей. Эта модель построена на определении сущностей (объектов) интереса пользователей библиографических записей, атрибутов каждой сущности и типов взаимосвязей между сущностями.

Для лучшего понимания механизмов построения библиографических связей стоит рассмотреть их основные виды. В формате *Universal Machine Readable Cataloguing* (далее – UNIMARC) используются горизонтальные, вертикальные и хронологические связи. Подробно проанализированы эти связи в основательном исследовании ведущего специалиста Библиотеки Конгресса США Б. Тиллет. Исследовательница идентифицировала и классифицировала библиографические связи, разделив их на семь основных категорий:

1) *эквивалентные* – устанавливаются между точными копиями одно-

го и того же произведения или между оригиналом и его репродукцией, при условии сохранения интеллектуального, художественного содержания и авторства;

2) *производные* – устанавливаются между библиографической единицей и модификацией, которая базируется на той же единице, включая варианты, версии, переводы, адаптации и т. п.;

3) *описательные* – устанавливаются между библиографической единицей и произведением, критической статьей, отзывом, рецензией, комментарием;

4) *вертикальные* – связи «целое-часть» или «часть-целое», которые устанавливаются между библиографической единицей и частью единицы (серией и выпуском серии);

5) *сопроводительные* – устанавливаются между библиографической единицей и приложением, дополнением;

6) *последовательные* – хронологические, устанавливаются между библиографическими единицами, одна из которых является продолжением другой или предшествует ей (например, продолжение издания объекта под другим названием);

7) *общие характеристики* – устанавливаются между библиографическими единицами, которые имеют общего автора, название или другие характеристики и являются дополнительными точками доступа в каталоге [22].

Библиотеки оказались в сложной, меняющейся информационной среде – электронные библиотечные ресурсы выросли в объемах и усложнились, изменились требования пользователей к информационному поиску. Цифровое пространство отличается многомерностью: различные пользователи, различные ресурсы, различные пути доступа к информации. 2015 г. стал определенным пределом окончательного понимания того, что библиотеки уже никогда не будут такими, какими были раньше, и эта мысль зафиксирована знаменательным событием – декларацией OCLC о прекращении печати каталожных карточек: «1 октября 2015 г. – после 44 лет и 1,9 миллиарда карточек – OCLC напечатал последнюю каталожную карточку» [25]. Изменения в семантике информационной среды существования библиотечных ресурсов, стандартах описания и обмена библиографическими данными повлияли на библиотечные электронные каталоги, что стало причиной появления понятия библиотечных каталогов следующего (нового) поколения (*Next-Generation Library Catalogs*).

С точки зрения пользователей, которые быстро привыкли к преимуществам поисковых систем Интернета, таких как Google, традиционная библиотечная система поиска оказалась громоздкой и неудобной. Насущной необходимостью стала интеграция имеющихся библиотечных электронных ресурсов и обеспечения читателей легкой, единственной и мощной платформой электронного ресурса. В 2007 г. Федерацией цифровых библиотек (*Digital Library Federation – DLF*) была создана специальная «Рабочая группа поисковых интерфейсов» (*Discovery Interfaces Task Force – DITF*), которая провела исследование потенциальных проблем интеграции библиотечной информационно-поисковой системы с внешними средствами веб-поиска. Результаты показали, что современные библиотечные поисковые системы (электронные каталоги – *online public access catalog*, далее – *OPAC*) имеют следующие проблемы:

- ориентированы на управление печатными книгами и полками в фонде (библиотечным собранием на бумажных носителях), во многих случаях не приспособлены к электронным ресурсам;
- основываются на библиографических стандартах данных, не поддерживают стандарты обмена данными с другими системами баз данных;
 - унифицированный библиографический поиск предназначен только для библиотечных ресурсов и не связан с веб-пространством;
 - интерфейс недостаточно дружелюбный – пользователи привыкли к простому и интуитивно понятному окну поиска браузера, поэтому они путаются и тяжело адаптируются к мощным функциям поиска библиотечной системы;
 - автоматически не исправляются дефекты представления поисковых терминов – не поддерживаются алгоритмы корреляции, не исправляется неправильное написание, не хватает семантической поддержки;
 - ответы на поисковый запрос очень сильно зависят от точности формулировки запроса пользователем (если пользователь не может предоставить точных библиографических данных, он не сможет получить желаемого результата) [19].

Исследователи выявили характерные черты поведения современных пользователей библиотечных электронных ресурсов, сформировавшиеся под влиянием удобных и гибких веб-технологий:

- задают неточные запросы длиной одно – три слова (ожидают, что компьютер угадает их потребности);

- хотят видеть контекст (все, что связано с запросом: связанные понятия и имеющиеся близкие по содержанию информационные ресурсы);
- проводят разведку по определенной теме, во многих случаях выбранной по Википедии (хотя просто нажать кнопку с названием темы, просмотреть списки документов, неоднократно вернуться, увидеть новые поступления и связанную с темой справочную информацию);
- не хотят думать (в эру алгоритмов, они не хотят выяснять, что им доступно и соответствует их потребностям, компьютер должен все это им предложить) [25, 27].

Понимание электронного каталога как ресурса, содержащего перечень библиографических записей, раскрывающих состав имеющихся в фонде библиотеки документов, уже не отвечает требованиям современных пользователей. Сейчас электронные каталоги содержат библиографические записи не только книг, но и статей, гипертекстовых документов, интернет-ресурсов. Такие БД могут раскрывать печатные и электронные ресурсы многих библиотек. Пользователи обращаются не к электронным каталогам, а к интегрированным библиотечно-информационным системам, имеющим расширенный перечень поисковых возможностей. Интегрированные библиотечно-информационные системы повышают полноту, оперативность библиотечно-информационного обслуживания всех категорий пользователей системы, обеспечивая свободный и широкий доступ к библиотечно-информационным ресурсам, позволяют сэкономить трудовые, финансовые ресурсы библиотек и информационных служб участников корпорации.

Тема интегрированного (федеративного, комбинированного) поиска информации начала интересовать библиотечных специалистов еще в 2001 г. В 2007 г. в издательстве Haworth Information Press (Binghamton, NY) вышел сборник статей по интегрированному поиску, который зарубежные исследователи обычно называют федеративным (*Federated Search*) [13]. Согласно онлайн-словарю по библиотечному делу и информатике (*Online Dictionary for Library and Information Science – ODLIS*) интегрированный (федеративный) поиск – это поиск информации с использованием специального программного обеспечения, предназначенного для направления запросов через единый интерфейс к нескольким разным сетевым информационным ресурсам (электронные каталоги, электронные библиотеки и научно-информационные базы данных) [23]. Интегрированный поиск делает документы, не всегда видимые поисковыми машинами Интернета («глубокий Веб»), доступными для поиска

без необходимости для пользователя посещать каждую БД отдельно. Для пользователей библиотек с развитой системой электронных каталогов, полнотекстовых БД, как собственных, так и внешних, возможность получить данные из множества источников по одному запросу является значительным преимуществом: экономит время, существенно совершенствует процесс сбора данных, обеспечивает охват более широкого диапазона информационных источников. Более развитые системы интегрированного поиска имеют функции ранжирования документов по релевантности и кластеризации результатов поиска.

В это же время, в 2007 г. М. Бридингом (M. Breeding) были сформулированы основные требования к современным библиотечным каталогам, которые должны соответствовать таким основным характеристикам:

- единая точка доступа ко всем имеющимся в библиотечном фонде ресурсам;
- высокий уровень дизайна и функциональности (как у коммерческих сайтов);
- расширение библиотечной информации за счет данных из внешних источников (обложки изданий, аннотации, рейтинги и т.д.);
- фасетный характер навигации в каталоге, возможность уточнять запрос по различным параметрам (дата, язык текста, доступа, предметная рубрика и т. д.);
- ранжированный характер вывода документов на экран (например, по частоте книговыдачи), возможность сортировки выборки по году издания, названию, автору и тому подобное;
- возможность поиска в режиме простого поискового окна (a la Google), возможность перехода к расширенному поиску;
- возможность проверки правильности написания слов;
- подсказки рекомендации дополнительных источников информации («смотри также по этой теме», «читатели вместе с этой книгой заказывают также ...»);
- пользователь может самостоятельно (с последующей модерацией специалиста) оставлять тематические теги и комментарии к записям каталога;
- связь с пользователями через RSS с возможностью получать информацию о новых поступлениях [11].

После того, как сервисы поиска библиотечных ресурсов успешно интегрировали информационные ресурсы библиотеки, пользователи

начали повышать требования к системе поиска знаний. Как отмечают исследователи, самой популярной системой содержательной навигации оказалась фасетная, большинство пользователей охотно ограничивают область поиска с помощью заданного множества параметров и гораздо меньше обращаются к специальным поисковым библиотечным инструментам: классификаторам, рубрикаторам и авторитетным файлам. Для успеха у пользователей важно использование этого библиотечного инструментария при поиске в скрытом («латентном») виде как простых интерактивных подсказок к найденным информационным массивам [27]. Первым комплексным изложением вопросов фасетной навигации стал курс лекций Д. Танкеланга (D. Tunkelang) «Фасетный поиск», изданный в 2009 г. издательством «Morgan & Claypool Publishers». Как отмечает автор лекций, фасетный поиск сразу показывает пользователю очевидную информационную структуру, передвигаясь по которой он видит наполнение информационного массива, к которому он обращается. В отличие от привычного параметрического поиска, где пользователь сразу должен точно задать все поисковые параметры в одном запросе, фасетная навигация позволяет постепенно уточнять запрос, интерактивно выбирая дополнительные параметры. Фасетная навигация также дает возможность сочетать в структурированные множества информацию, совместить которую в обычных запросах по ключевым словам чрезвычайно сложно. Кроме того, визуальный интерактивный интерфейс фасетной навигации предотвращает возможные ошибки и отказы, он намного проще и привлекательнее, чем текстовый запрос из командной строки [28]. Все эти преимущества фасетной навигации сделали ее одним из важных элементов инструментария усовершенствованных систем выявления ресурсов и знаний (*Discovery System*).

Начиная с 2016 г. дискуссия по поводу того, какими должны быть онлайн-библиотечные каталоги, переместилась от обсуждения каталогов нового поколения в сторону обсуждения того, что библиотечный поисковый сервис должен быть основан на системе выявления знаний (*Knowledge Discovery System* или *Discovery System*), схожей с сервисами поисковых систем Интернета.

Изменения в терминологии библиотечного поиска анализирует американская исследовательница И. Трапидо (*I. Trapido*) [27], и в этом контексте она обращает внимание на такие ключевые терминологические конструкции:

- *электронный каталог (OPAC – online public access catalog) –*

онлайн-каталог библиотеки, ориентированный на заказ литературы из библиотечного фонда;

- *системы выявления ресурсов (discovery system)* – интерфейсы каталогов нового поколения, отделенные от интегрированной библиотечной системы, которые предлагают расширенные функциональные возможности, такие как фасетная навигация, ранжированные результаты поиска, а также возможность включения в результаты поиска содержания институциональных репозитариев и электронных библиотек;

- *веб-ориентированные инструменты поиска (web-scale discovery tools)* – системы, которые дополнительно к функциональности каталогов следующего поколения расширяют сферу поиска путем систематической агрегации содержания библиотечных каталогов, подписных БД и институциональных репозитариев в единый центральный индекс, который также попадает в индекс поисковых систем Интернета.

Обобщенно исследователи отмечают следующие характеристики библиотечных поисковых интерфейсов, основанных на технологии выявления ресурсов:

- *объем*: по сравнению с ОПАС поиск не ограничивается одной базой данных библиотечного каталога, объем ее информационного охвата является широким и гибким, объединяет все библиотечные ресурсы;

- *содержание*: сочетает традиционный ОПАС и базу данных электронных ресурсов в едином индексе;

- *поиск*: при поиске в различных базах данных есть возможность быстро и оперативно получить рекомендации и различные релевантные информационные результаты, которые лучше всего отвечают поисковому запросу, независимо от местонахождения и формата ресурсов;

- *доступ*: поддерживается функция простого поиска и фасетная навигация, имеются рекомендации, связанные с содержанием строки поиска. С точки зрения пользователей, ключевым элементом является возможность чувствовать, что процесс от нахождения до получения информации проходит быстро, беспрепятственно и просто [29].

Исполнительный директор службы OCLC, отвечающий за конечного пользователя, М. Шоултер (*M. Showalter*) в своем докладе в 2015 г. подытожил основные привлекательные для пользователей характеристики современного интерфейса выявления библиотечных ресурсов:

- поиск всего контента в одном месте;
- центральный индекс локального фонда + электронные ресурсы;
- простой быстрый поиск по ключевым словам;

- наличие заданных фасет для фильтрации результатов поиска;
- современный, интуитивный интерфейс;
- получение полного текста одним нажатием («*one-click to full text*»)

[25].

Сейчас есть много систем, в которых реализованы основные особенности электронных каталогов нового поколения: коммерческие (Aquabrowser, BiblioCommons, EBSCO Discovery Service, Endeca) и системы с открытым кодом (Blacklight, Fac-Back-OPAC, LibraryFind, Rapi) [26]. Веб-ориентированный инструментарий, обеспечивающий оперативный доступ к библиографическим и полнотекстовым ресурсам библиотек мира, реализован в таких проектах: Всемирный каталог – WorldCat (<https://www.worldcat.org/>), Всемирная цифровая библиотека – World Digital Library (<https://www.wdl.org/>), проект Европейской электронной библиотеки Европеана – Europeana (<http://www.europeana.eu>). Для академической литературы платформы выявления ресурсов, такие как Web of Science (<https://www.webofknowledge.com>) и Google Scholar (<https://scholar.google.com>), предлагают простой метод поиска широкого диапазона научных источников по многим дисциплинам: академическая пресса, профессиональные форумы, статьи, академические отчеты университетов и научно-исследовательских институтов, дипломные работы, книги, рефераты и др.

Среди библиотечных проектов, реализующих принципы систем выявления знаний и открытых связанных данных, следует отметить проект Национальной библиотеки Франции (далее – НБФ) data.bnf.fr (<http://data.bnf.fr/>), целью которого является: сделать библиотечные данные более видимыми в Интернете, объединить библиотечные данные, как в каталогах, так и вне их; способствовать созданию связей между структурированными и доверенными веб-ресурсами, обеспечить повторное использование метаданных (под открытой лицензией) третьей стороной. К важным библиотечным данным относятся названия произведений, персоналии и организации. Страницы структурированных данных наделяют определенное зафиксированное понятие идентификатором, собирают доступную информацию о нем из разных справочных веб-источников, обеспечивают пользователю удобный переход к соответствующим информационным источникам из фондов НБФ (цифровая библиотека Gallica, генеральный электронный каталог, каталог архивных и рукописных документов) и внешних библиотечных и информационных проектов WorldCat, Europeana, Wikipedia. Такие

страницы данных хорошо индексируются поисковыми работами и появляются в первых строках ответов на информационный запрос, что значительно улучшает видимость и доступность библиотечных ресурсов в Интернете. Между данными НБФ и WikiData обеспечена связь, которая позволяет актуализировать и пополнять библиотечные данные, в частности, из Wikipedia автоматически выводятся изображения и краткие аннотации для соответствующих данных НБФ [7]. НБУВ представлена в data.bnf.fr статьей «Bibliothèque nationale VI Vernadsky d'Ukraine», из которой можно получить информацию о том, что библиотека находится в Украине, основана в 1918 г., в фонде НБФ есть пять изданий национальной библиографии, изданных НБУВ, есть соответствующая запись в VIAF.

Эти современные подходы организации доступа к библиотечным информационным ресурсам побудили инициировать в НБУВ внедрение семантических структур связанных данных и инструментов фасетной навигации в рамках двух информационных проектов: «Наука Украины: доступ к знаниям» и электронная библиотека «Украиника».

Портал «Наука Украины: доступ к знаниям» содержит структурированные данные о научных учреждениях, библиотеках и их ресурсах, ученых Украины. Авторские поисковые профили используются для выявления библиографической информации в различных научных информационных ресурсах НБУВ: электронном каталоге, электронной библиотеке «Научная периодика Украины», реферативной базе данных. Пользователь также информируется о наличии соответствующих произведений в электронной библиотеке «Украиника», архивном и рукописном фонде НБУВ. Обеспечена возможность поиска коллег по научной специальности и соавторов произведений. Страница авторского поискового профиля имеет уникальный идентификатор и ссылки к авторским страницам в других информационных системах (ORCID, Google Scholar, ResearcherID, Scopus, Wikipedia и т.п.) [2]. В 2020 г. стартовал новый сервис научных метаданных ORCID и DOI для ученых Украины. В профилях ученых дополнительно к информации о научных e-ресурсах НБУВ динамически отображаются публикации, зарегистрированные международной научной базой данных CrossRef (<http://search.crossref.org>), которая содержит метаданные более 100 млн научных публикаций. Список публикаций автоматически формируется по авторскому идентификатору – ORCID, указанному в профиле ученого на портале НБУВ. Каждая библиографическая ссылка в списке содержит DOI

публикации (по которому можно перейти на оригинальную страницу публикации).

Электронную библиотеку «Украиника» сопровождает уникальный справочный аппарат структурированных связанных данных, относящихся к информационному полю знаний об Украине. Каждая страница данных содержит уникальный идентификатор, название, информационное изображение, категорию, варианты написания заголовка, информацию о связанных данных, гиперссылки к справочным интернет-ресурсам. В электронной библиотеке выделены различные категории предметных справок, которые организуют фасетную навигацию в массиве документов: «Персоналии», «Исторические события», «Слой населения», «Географические объекты», «Государства», «Народы», «Населенные пункты», «Памятники культуры», «Памятники природы», «Учреждения» и другие. Каждая категория сопровождается своим типом изображений (портрет персоналии, картина или фото, хрестоматийное изображение, типичная фигура, здание, флаг, герб, логотип, народный костюм, карта, буквенный акроним и т. д.). Дополнительно к текстовой информации справочник хранит и систематизирует визуальный ряд, связанный со знаковыми понятиями украиники. Для каждой справки автоматически появляется информация о наличии соответствующих произведений в электронной библиотеке, для каждого произведения – информация об имеющихся справочных ресурсах. Для автора-ученого приводится ссылка к его авторскому профилю на портале «Наука Украины: доступ к знаниям» [4]. В дальнейшем предполагается провести интеграцию справочного материала «Украиники» с данными WikiData, что позволит автоматически получать дополнительную информацию из открытых источников данных.

Внедрение фасетного поиска и системы связанных данных для ключевых поисковых сервисов НБУВ обеспечивает постепенное расширение онтологической базы знаний об Украине и ее науке, совершенствует доступ к национальным и научным ресурсам НБУВ, способствует видимости электронных ресурсов НБУВ поисковыми системами Интернета. Параллельно с развитием справочного аппарата связанных данных в НБУВ также проводятся работы по созданию национальных авторитетных файлов лиц и учреждений. Авторитетные файлы сопровождается информация об имеющихся идентификаторах соответствующих данных в международных системах авторитетного нормативного контроля: WorldCat, ISNI, VIAF, WikiData. Все эти работы обеспечат активное

вхождение информационных ресурсов НБУВ в систему семантического поиска современной веб-среды.

Список использованных источников

1. Горшков С. Введение в онтологическое моделирование: ревизия 2.3. ООО «ТриниДата». 2014–2018. URL: <https://trinidata.ru/files/SemanticIntro.pdf>.

2. Лобузина Е. В. Информационный портал «Наука Украины: доступ к знаниям». *Библиотеки национальных академий наук: проблемы функционирования, тенденции развития*. 2017. Вып. 14. С. 35–46. URL: http://nbuv.gov.ua/UJRN/bnan_2017_14_6.

3. Лобузина К. В. Библиотека 3:0: знания, сховища даних та експерти. *Бібліотекознавство. Документознавство. Інформологія*. 2012. № 1. С. 26–35. URL: http://nbuv.gov.ua/UJRN/bdi_2012_1_7.

4. Лобузина К., Перенесієнко І., Вощенко О. Система організації тематичних зв'язків предметно-довідкового апарату електронної бібліотеки «Україніка». *Наукові праці Національної бібліотеки України імені В. І. Вернадського*. Київ, 2019. Вип. 55. С. 220–230.

5. Серебряков В. А. Что такое семантическая цифровая библиотека. *Труды 16-й Всероссийской научной конференции «Электронные библиотеки: перспективные методы и технологии, электронные коллекции» – RCDL-2014* (Дубна, 13–16 октября 2014). Дубна: ОИЯИ, 2014. С. 21–25.

6. Belling A., Rhodes A., Smith J., Thomson S., Thorn B. Exploring Library 3.0 and beyond. 2010–2011 Shared Leadership Program State Library of Victoria and Public Libraries Victoria Network. 2011. URL: http://www.libraries.vic.gov.au/downloads/20102011_Shared_Leadership_Program_Presentation_Day_/exploring_library_3.pdf.

7. Bernes E., Boulet V., Leclair C. Améliorer l'accès aux données des bibliothèques sur le web : l'exemple de data.bnf.fr. *IFLA World Library and Information Congress*. 2016. URL: <http://library.ifla.org/1447/1/081-bernes-fr.pdf>.

8. Berners-Lee T., Hendler J., Lassila O. The Semantic Web : A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. *Scientific American*. 2001. URL: <http://www.sciam.com> (May 17, 2001).

9. Berners-Lee T. Linked Data. *W3C*. 2006. URL: <https://www.w3.org/DesignIssues/LinkedData.html> (last change: 18.06.2009).

10. Bibliographic Framework Initiative. *Library of Congress, USA*. URL: <https://www.loc.gov/bibframe/>.

11. Breeding M. Next-Generation Library Catalogs. *Library Technology Reports. Chapter 1: Introduction*. 2007. Vol. 43, No 4., pp. 5–14. URL: http://alatechsource.metapress.com/content/p6_rll2_h042601168/fulltext.pdf.

12. Expression of Core FRBR Concepts in RDF. *Vocab.org – A URI space for vocabularies*. 2005. URL: <http://vocab.org/frbr/core.html>.

13. Federated Search: Solution or Setback for Online Library Services / edited by Christopher N. Cox. Binghamton, NY: Haworth Pressm, 2007. 374 p.

14. Functional Requirements for Bibliographic Records – FRBR. 1999. URL: <http://www.ifla.org/publications/functional-requirements-for-bibliographic-records>

15. G8 Open Data Charter. URL: <https://opendatacharter.net/g8-open-data-charter/>.

16. Gartner Names Semantic Technologies To Its Top Technology Trends Impacting Information Infrastructure in 2013. DATAVERSITY. 2013. URL: <https://www.dataversity.net/gartner-names-semantic-technologies-to-its-top-technology-trends-impacting-information-infrastructure-in-2013/>.

17. Gartner Top 10 Strategic Technology Trends for 2019. *Smarter with Gartner*. 2018. URL: <https://www.gartner.com/smarterwithgartner/gartner-top-10-strategic-technology-trends-for-2019/>.

18. Gucer V. 5 Things To Know About Semantic Technologies. *IBM Redbooks Blog*. 2013. URL: https://www.ibm.com/developerworks/community/blogs/5_things/entry/5_things_to_know_about_the_semantictchnologies?lang=en.

19. ILS–DI [Integrated Library System – Discovery Interface Task Group]. *Digital Library Federation*. URL: <http://www.diglib.org/project-update-ils-di/>.

20. Kwanya T., Stilwell C., Underwood P. *Library 3.0: Intelligent Libraries and Apomediation*. Chandos Publishing, 2014. 190 p.

21. Linked Data – Connect Distributed Data across the Web. URL: <http://linkeddata.org/>.

22. Noruzi A. FRBR and Tillett’s Taxonomy of Bibliographic Relationships. *Knowledge Organization*. Vol. 39, is. 6. 2012. pp. 409–416. URL: <http://eprints.rclis.org/25682/>.

23. Online Dictionary for Library and Information Science – ODLIS. *ABC–CLIO Corporate*. URL: https://www.abc-clio.com/ODLIS/odlis_about.aspx.

24. Saw G., Todd H. Library 3.0: where are our skills? *World library and information congress : 73 rd IFLA general conf. and council (19–23 Aug. 2007, Durban, South Africa)*. Durban, 2007. pp. 1–15.

25. Showalter M. Library Discovery: Challenges and Opportunities. *OCLC EMEA Regional Council 2016 (EMEARC16), 1–2 March, Madrid. 2016*. URL: https://www.oclc.org/content/dam/oclc/events/2016/EMEARC2016/EMEARC-16_Session-H_WorldCat-Discovery-Services.pdf.

26. Stevenson K., Elsegood S., Seaman D. [et al.] Next-generation library catalogues: reviews of Encore, Primo, Summon and Summa. *Serials*. 2009. 22 (1). pp. 68–82.

27. Trapido I. Library Discovery Products: Discovering User Expectations through Failure Analysis. *Information Technology and Libraries*. 2016. 35(3). Pp. 9–26. <https://doi.org/10.6017/ital.v35i3.9190>.

28. Tunkelang D. Faceted search. Synthesis lectures on information concepts, retrieval, and services, Lecture #5. Morgan & Claypool Publishers, 2009. 96 p. <https://doi.org/10.2200/S00190ED1V01Y200904ICR005>.

29. Wang Z. Analysis of Digital Library Resource Discovery Service System. *International Conference on Education, Management, Information and Medicine (EMIM 2015), April 24–26, 2015*. Shenyang, China. 2015. DOI: <https://doi.org/10.2991/emim-15.2015.142>.

30. WikiData. URL: <https://www.wikidata.org/>.

References

1. Gorshkov, S. (2014–2018). Vvedeniie v ontologicheskoiie modelirovaniie: reviziia 2.3 [Introduction to ontological modeling: revision 2.3]. OOO «TriniData». Retrieved from <https://trinidata.ru/files/SemanticIntro.pdf> [in Russian].

2. Lobuzina, E. V. (2017). Informatsionnyi portal «Nauka Ukrainy: dostup k znaniiam» [Information portal «Science of Ukraine: access to knowledge»]. *Biblioteki natsionalnykh akademii nauk : problemy funktsionirovaniia, tendentsii razvitiia – Libraries of national academies of sciences: problems of functioning tendencies of development*, 14, 35–46. Retrieved from http://nbuv.gov.ua/UJRN/bnan_2017_14_6 [in Russian].

3. Lobuzina, K. V. (2012). Biblioteka 3:0: znannia, skhovyshcha danykh ta eksperty [Library 3: 0: knowledge, data warehouses and experts]. *Bibliotekoznavstvo. Dokumentoznavstvo. Informolohiia – Library science*.

Documentation. Informology, 1, 26–35. Retrieved from http://nbuv.gov.ua/UJRN/bdi_2012_1_7 [in Ukrainian].

4. Lobuzina, K., Perenesiienko, I., Voshchenko, O. (2019). Systema orghanizatsii tematychnykh zviazkiv predmetno-dovidkovogho aparatu elektronnoi biblioteky «Ukrainika» [The system of organization of thematic relations of the subject-reference apparatus of the electronic library «Ukrainica»]. *Naukovi pratsi Natsionalnoi biblioteky Ukrainy im. V. I. Vernadskoho – Transactios of V. I. Vernadsky National Library of Ukraine*, 55, 220–230. Retrieved from http://np.nbuv.gov.ua/doc/npnbuimviv_2019_55_17 [in Ukrainian].

5. Serebriakov, V. A. (2014). Chto takoe semanticheskaia tsifrovaia biblioteka [What is a semantic digital library]. *Proceedings of the Digital libraries: advanced methods and technologies, digital collections (RCDL–2014): 16-ia Vserossiiskaia nauchnaia konferentsiia (Dubna, 13–16 oktiabria 2014 h.) – 16 th All-Russian Scientific Conference* (pp. 21–25). Dubna [in Russian].

6. Belling, A., Rhodes, A., Smith J., Thomson S., Thorn B. (2011). Exploring Library 3.0 and beyond. 2010–2011 Shared Leadership Program State Library of Victoria and Public Libraries Victoria Network. Retrieved from http://www.libraries.vic.gov.au/downloads/20102011_Shared_Leadership_Program_Presentation_Day_/exploring_library_3.pdf [in English].

7. Bermes, E., Boulet, V., Leclair, C. (2016) Améliorer l'accès aux données des bibliothèques sur le web : l'exemple de data.bnf.fr. IFLA World Library and Information Congress. Retrieved from <http://library.ifla.org/1447/1/081-bermes-fr.pdf> [in French].

8. Berners-Lee, T., Hendler, J., Lassila, O. (2001) The Semantic Web : A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. *Scientific American*. Retrieved from <http://www.sciam.com> [in English].

9. Berners-Lee, T. (2006) Linked Data. *W3C*. Retrieved from <https://www.w3.org/DesignIssues/LinkedData.html> (last change: 18.06.2009) [in English].

10. Bibliographic Framework Initiative. *Library of Congress, USA*. Retrieved from <https://www.loc.gov/bibframe/> [in English].

11. Breeding, M. (2007) Next-Generation Library Catalogs. *Library Technology Reports. Chapter 1: Introduction*, 43 (4), 5–14. Retrieved from http://alatechsource.metapress.com/content/p6_rll2_h042601168/fulltext.pdf [in English].

12. Expression of Core FRBR Concepts in RDF (2005). *Vocab.org* –

A URI space for vocabularies. Retrieved from <http://vocab.org/frbr/core.html> [in English].

13. *Federated Search: Solution or Setback for Online Library Services (2007)* / edited by Christopher N. Cox. Binghamton, NY: Haworth Press. 374 p. [in English].

14. *Functional Requirements for Bibliographic Records – FRBR (1999)*. Retrieved from <http://www.ifla.org/publications/functional-requirements-for-bibliographic-records> [in English].

15. *G8 Open Data Charter*. Retrieved from <https://opendatacharter.net/g8-open-data-charter/> [in English].

16. *Gartner Names Semantic Technologies To Its Top Technology Trends Impacting Information Infrastructure in 2013 (2013)*. *DATAVERSITY*. Retrieved from <https://www.dataversity.net/gartner-names-semantic-technologies-to-its-top-technology-trends-impacting-information-infrastructure-in-2013/> [in English].

17. *Gartner Top 10 Strategic Technology Trends for 2019 (2018)*. *Smarter with Gartner*. Retrieved from <https://www.gartner.com/smarterwithgartner/gartner-top-10-strategic-technology-trends-for-2019/> [in English].

18. Gucer, V. (2013) *5 Things To Know About Semantic Technologies*. *IBM Redbooks Blog*. Retrieved from https://www.ibm.com/developerworks/community/blogs/5_things_to_know_about_the_semantic_technologies?lang=en [in English].

19. ILS–DI [Integrated Library System – Discovery Interface Task Group]. *Digital Library Federation*. Retrieved from <http://www.diglib.org/project-update-ils-di/> [in English].

20. Kwanya, T., Stilwell, C., Underwood, P. (2014) *Library 3.0: Intelligent Libraries and Apomediation*. Chandos Publishing. 190 p. [in English].

21. *Linked Data – Connect Distributed Data across the Web*. Retrieved from <http://linkeddata.org/> [in English].

22. Noruzi, A. (2012) *FRBR and Tillett’s Taxonomy of Bibliographic Relationships*. *Knowledge Organization*, 39 (6), 409–416. Retrieved from <http://eprints.rclis.org/25682/> [in English].

23. *Online Dictionary for Library and Information Science – ODLIS*. *ABC–CLIO Corporate*. Retrieved from https://www.abc-clio.com/ODLIS/odlis_about.aspx [in English].

24. Saw, G., Todd, H. (2007) *Library 3.0: where art our skills? World library and information congress : 73 rd IFLA general conf. and council (19–23 Aug. 2007, Durban, South Africa)*. Durban. pp. 1–15 [in English].

25. Showalter, M. (2016) Library Discovery: Challenges and Opportunities. *OCLC EMEA Regional Council 2016 (EMEARC16), 1–2 March, Madrid*. Retrieved from https://www.oclc.org/content/dam/oclc/events/2016/EMEARC2016/EMEARC-16_Session-H_WorldCat-Discovery-Services.pdf [in English].

26. Stevenson, K., Elsegood, S., Seaman, D. [et al.] (2009) Next-generation library catalogues: reviews of Encore, Primo, Summon and Summa. *Serials*. 22 (1), 68–82 [in English].

27. Trapido, I. (2016). Library Discovery Products: Discovering User Expectations through Failure Analysis. *Information Technology and Libraries*, 35(3), 9–26. <https://doi.org/10.6017/ital.v35i3.9190> [in English].

28. Tunkelang, D. (2009). Faceted search. Synthesis lectures on information concepts, retrieval, and services, Lecture #5. Morgan & Claypool Publishers. 96 p. <https://doi.org/10.2200/S00190ED1V01Y200904ICR005> [in English].

29. Wang Z. (2015) Analysis of Digital Library Resource Discovery Service System. *International Conference on Education, Management, Information and Medicine (EMIM 2015), April 24–26, 2015, Shenyang, China*. DOI: <https://doi.org/10.2991/emim-15.2015.142> [in English].

30. WikiData. Retrieved from <https://www.wikidata.org/> [in English].

Катерина Лобузина,

директор Інституту інформаційних технологій
Національної бібліотеки України імені В. І. Вернадського,
доктор наук із соціальних комунікацій,
старший науковий співробітник,
Україна, м. Київ

e-mail: lobuzina@nbuv.gov.ua

ORCID: 0000-0003-3371-4029

СЕМАНТИЧНІ ВЕБТЕХНОЛОГІЇ ТА БІБЛІОТЕЧНІ ПОШУКОВІ СИСТЕМИ

Статтю присвячено використанню семантичних вебтехнологій у бібліотечно-інформаційній діяльності. Розглянуто принципові зміни, що відбулися в сучасному вебсередовищі, їх вплив на еволюцію бібліотечних пошукових систем. Визначено поняття семантичної бібліотеки (Бібліотеки 3.0), яка використовує методи доступу до інформації, засновані на складних семантичних відносинах. Висвітлено зміни підходів до організації електронних бібліотечних ресурсів і систем пошуку інформаційних джерел. Показано еволюцію бібліо-

течних інформаційних систем від онлайн-електронних каталогів до систем виявлення знань. Проаналізовано основні характеристики, яким мають відповідати сучасні бібліотечні каталоги з точки зору інтернет-користувачів: пошук усього контенту в одному місці, ранжування виведення документів на екран, простий пошук за ключовими словами, фасети для фільтрації результатів пошуку, сучасний інтуїтивний інтерфейс, отримання повного тексту одним натисканням, підказки рекомендації додаткових джерел інформації, розширення бібліотечної інформації за рахунок даних із зовнішніх джерел. Охарактеризовані підходи використання технологій Семантичного Вебу в бібліотечних пошукових системах: федеративний пошук, пов'язані дані, фасетний пошук. Продемонстровані способи використання структурованих відкритих пов'язаних даних, таких як WikiData, в бібліотечних пошукових системах, що значно покращує видимість і доступність бібліотечних ресурсів в Інтернеті. Викладений матеріал проілюстровано досвідом впровадження семантичних вебтехнологій для пошуку інформації на порталі Національної бібліотеки України імені В. І. Вернадського.

Ключові слова: семантичні вебтехнології, інформаційний пошук, бібліотечні пошукові системи, електронні каталоги, бібліотечні каталоги нового покоління, системи виявлення знань, пов'язані дані.

Kateryna Lobuzina,

Head of Institut of Information Technologies

V. I. Vernadsky National Library of Ukraine,

PhD in Social Communications,

Ukraine, Kyiv

e-mail: lobuzina@nbuv.gov.ua

ORCID: 0000-0003-3371-4029

SEMANTIC WEB TECHNOLOGIES AND LIBRARY SEARCH SYSTEMS

The article focuses on the use of semantic web technologies in library and information activities. The fundamental changes that have taken place in the modern web environment, their influence on the evolution of library search engines are considered. The concept of a semantic library (Libraries 3.0) which uses methods of accessing information based on complex semantic relations, is defined. Changes in approaches to the organization of electronic library resources and search systems for information sources, are highlighted. The evolution of library information systems from online public access catalogs to knowledge discovery systems is shown. The main characteristics that modern library catalogs must meet from the point of view of Internet users,

have been analyzed: search for all content in one place, ranking of document display, simple keyword search, facets for filtering search results, modern intuitive interface, getting the full text with one click, recommendations for additional information sources, expanding library information with data from external sources. The approaches to the use of Semantic Web technologies in library search engines are characterized: federated search, linked data, faceted search. Ways to use structured open linked data such as WikiData in library search engines, greatly improving the visibility and accessibility of library resources on the Internet are demonstrated. The presented material is illustrated by the experience of introducing semantic web technologies for information retrieval on the portal of Vernadsky National Library of Ukraine.

Keywords: semantic web technologies, information retrieval, library search engines, electronic catalogs, new generation library catalogs, knowledge identification systems, linked data.