ТЕМАТИЧЕСКИЙ ПОИСК В ЭЛЕКТРОННОМ КАТАЛОГЕ





Наталия Зорина,
Научный сотрудник отдела комплектования и научной обработки документов Государственной научно-педагогической библиотеки Украины имени В.А. Сухомлинского, г. Киев

Благодаря использованию автоматизированных информационно-поисковых систем, формирование электронных каталогов и библиографических баз данных, обеспечение информационных потребностей пользователей осуществляется на качественно новом уровне

Активное формирование и развитие полноценного информационного ресурса, научно-информационное обеспечение инновационного развития образования, науки, профессиональных потребностей ученых и практиков системы образования Украины – важные задания современной научной библиотеки, каковой является Государственная научно-педагогическая библиотека Украины имени В.А. Сухомлинского (далее – ГНПБУ).

Внедрение в библиотечные процессы компьютерных технологий затронуло прежде всего процессы каталогизации. Благодаря использованию автоматизированных информационно-поисковых систем, формирование электронных каталогов и библиографических баз данных, обеспечение информационных потребностей пользователей осуществляется на качественно новом уровне. Возможность оперативного многоаспектного информационного поиска значительно расширяет его границы по сравнению с возможностями традиционных карточных каталогов, поскольку электронный каталог в целом охватывает всю их совокупность.

В условиях перехода библиотек к автоматизированному способу формирования библиографической записи ее качество определяется не только соблюдением государственных стандартов в процессе ее создания, но и прежде всего возможностью эффективного поиска документов, осуществление которого возможно лишь при условии четкой организации хранения документов и сведений о них, то есть соблюдение определенного порядка размещения документов в фондохранилище и их библиографических записей в поисковых массивах через группировку по определенным признакам (фамилия автора, название документа, год издания, вид документа, формат, носитель информации). Но в тематическом поиске, то есть в поиске по содержанию, именно этот признак является основным. Мнения специалистов относительно того, какой информационно-поисковый язык предпочитают пользователи для осуществления поиска в электронных каталогах, расходятся. Практический опыт показывает, что большинство читательских запросов имеют тематический характер, то есть пользователи предпочитают поиск по языкам вербального типа, к которым относится и язык предметных рубрик. К сожалению, поиск по предметным рубрикам пока не является приоритетным для значительной части пользователей, они преимущественно используют язык ключевых слов из-за его простоты и сходства с поиском в Интернете. Однако язык ключевых слов имеет ряд недостатков, обусловленных его ненормированностью. В то же время преимущество поиска по предметным рубрикам заключается в том, что каждая из них является законченной мыслью и дает точную информацию о документе, а комплекс предметных рубрик со всеми подзаголовками позволяет раскрыть большинство аспектов предмета. В случае, когда название документа неинформативно, именно предметные рубрики наиболее полно раскрывают содержание документа. Поэтому в наше время, учитывая, что в подавляющем большинстве научных библиотек электронный каталог стал основным инструментом доступа к информационным ресурсам, возрастает актуальность именно предметного ввода и поиска информации в электронном каталоге, что обусловливает необходимость смысловой обработки документов, внедрение и использование информационно-поискового языка предметного типа.

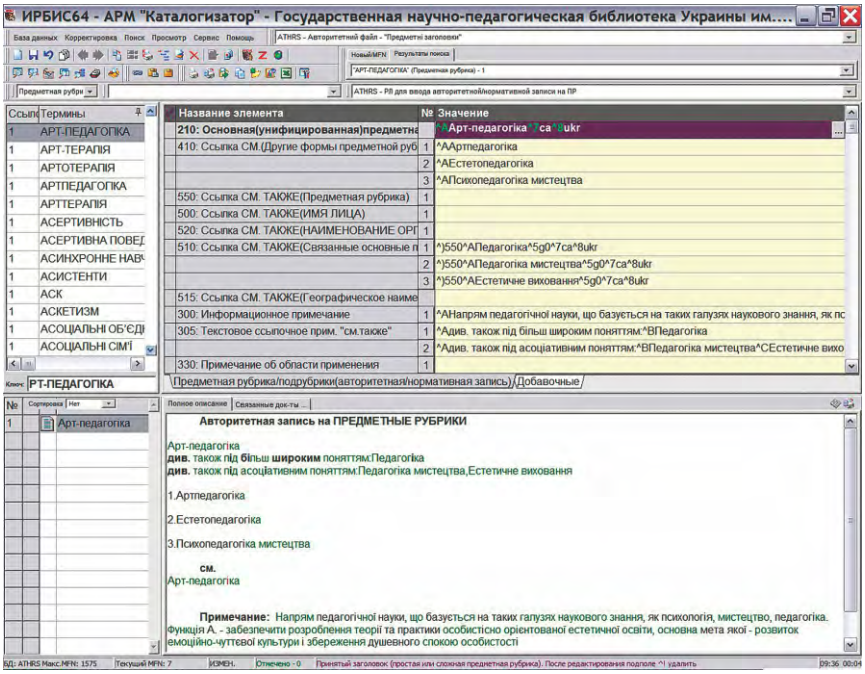
Поиск в электронных каталогах привлекает читателя прежде всего минимальными затратами времени. А возможность круглосуточного доступа к электронным каталогам через

Интернет позволяет работать с ними где угодно и в любое удобное время. Однако при осуществлении тематического поиска в электронном каталоге пользователи сталкиваются с определенными проблемами. В частности, результат поиска во многом зависит от формулировки запроса, следовательно читатель получает информацию в строгом соответствии со своим конкретным запросом. Здесь действуют два важных фактора: во-первых, четкое определение задачи, для выполнения которого осуществляют поиск, во-вторых, профессиональные и личностные характеристики пользователя, степень его подготовленности и владения научной терминологией. Если сформулированный запрос лексически не совпадает с имеющимися в электронном каталоге формулировками предметных рубрик, результат поиска может оказаться неудовлетворительным. Поэтому применение предметизационного информационно-поискового языка требует как от каталогизатора, так и от пользователя глубокого знания научной терминологии, а языковые средства электронного каталога, с одной стороны, должны иметь достаточную функциональную полноту для обеспечения выполнения различных поисковых задач, а с другой – должны быть адаптированы и для неподготовленных пользователей.

Предметизация как вид индексирования документов базируется на предметной классификации документов и предусматривает отражение содержания документа в понятийной системе естественного языка в виде предметных рубрик. Основной задачей предметизации является определение предмета как основной темы рассмотрения, а также выяснение основных аспектов характеристики предмета, его отношений и связей с другими предметами и формулирование содержания документа в словесной форме в виде предметных рубрик с соответствующей полнотой и точностью. Грамотно составленная предметная рубрика должна обеспечить адекватное отображение содержания документа, дать более точный результат поиска, чем поиск по ключевым словам и минимум информационного шума в процессе поиска нужной темы.

В качестве лексических единиц в предметизации используют слова и словосочетания естественного языка, что позволяет организовывать поиск в списке предметных рубрик по алфавиту. Сейчас особенно широко предметную классификацию используют в создании электронных каталогов, баз данных, тезаурусов, рубрикаторов, словарей и авторитетных нормативных файлов предметных рубрик.

ПРИМЕР АВТОРИТЕТНОЙ ЗАПИСИ В БАЗЕ ДАННЫХ АВТОРИТЕТНЫЙ ФАЙЛ «ПРЕДМЕТНЫЕ ЗАГОЛОВКИ»



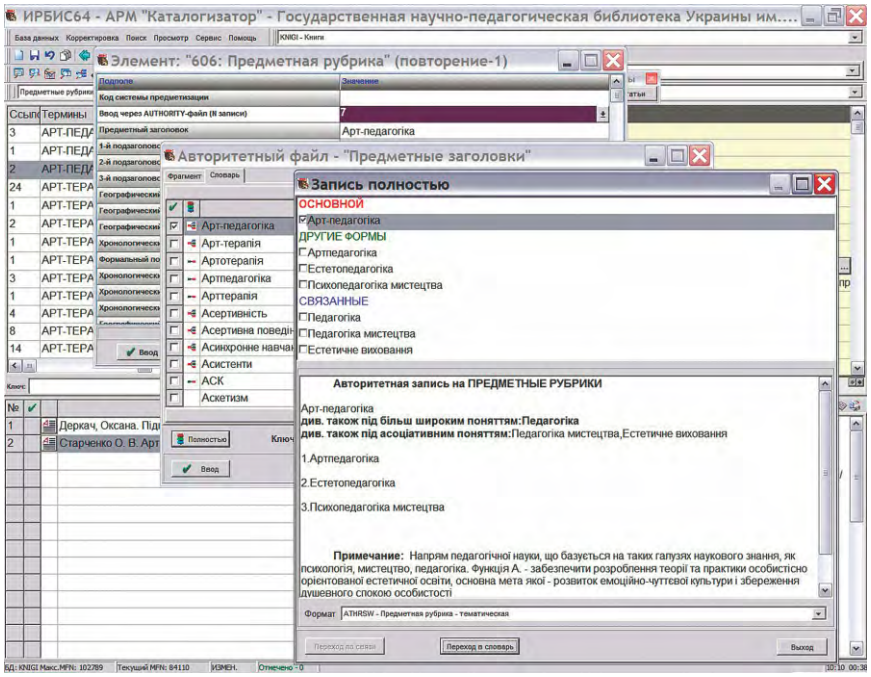
Статья посвящена рассмотрению проблемы унификации в процессе предметизации документов, особенностей формирования отраслевой базы данных Авторитетного файла «Предметные заголовки» по вопросам образования, педагогики и психологии и организации тематического поиска с ее использованием, в частности в ГНПБУ.

С внедрением системы электронного заказа документов, увеличения доли полнотекстовых электронных ресурсов электронный каталог все больше ориентируется на удаленного пользователя, который не может воспользоваться помощью библиотекаря-консультанта или традиционной системой каталогов и картотек. Поэтому поисковая навигация электронного каталога должна быть максимально удобной для самостоятельной работы и отвечать научным критериям.

В условиях электронной каталогизации весьма ощутима потребность в наличии полноценного списка заголовков предметных рубрик, наличие которого поможет оптимизировать как процесс предметизации, так и решить определенные проблемы поиска документов, в частности, избавит пользователя от формулирования запроса вслепую. Следует отметить, что большинство компьютерных

Предметизация как вид индексирования документов базируется на предметной классификации документов и предусматривает отражение содержания документа в понятийной системе естественного языка в виде предметных рубрик

ИСПОЛЬЗОВАНИЕ АВТОРИТЕТНОГО ФАЙЛА «ПРЕДМЕТНЫЕ ЗАГОЛОВКИ» В ПРОЦЕССЕ ПРЕДМЕТИЗАЦИИ ДОКУМЕНТОВ



Авторитетные файлы предметных рубрик обеспечивают возможность глубокого и результативного поиска в электронном каталоге, гарантируют его полноту и точность, служат контрольным словарем в организации тематического доступа к библиографическим записям

программ, обеспечивающих электронную каталогизацию, дают возможность просмотреть список предметных рубрик по вертикали и горизонтали. Вертикальный просмотр списка предметных рубрик по алфавиту расширяет границы предметного поиска по запросу благодаря подбору родственных предметных рубрик, например, выраженных словами, имеющими общий корень. Горизонтальный просмотр списка делает возможным уточнение, а в случае необходимости – расширение или сужение границ предметного поиска по запросу благодаря подрубрикам. Кроме того, система ссылок между предметными рубриками также делает расширение, сужение и уточнение границ предметного поиска и непосредственно влияет на его качество.

Для межбиблиотечного сотрудничества наличие такого списка также является чрезвычайно важным фактором, поскольку с появлением электронных каталогов предметные файлы библиотек разного типа и профиля перестают быть замкнутыми системами, обслуживающих только определенный круг пользователей, и вопрос унификации предметизации является сегодня особенно актуальным, как в отношении возможности однократной обработки документов, так и широкого обмена информацией. Важную организационную роль в этом играет система машиночитаемых форматов каталогизации.

Внедрение и развитие различных форм межбиблиотечного взаимодействия, таких, как создание сводных электронных каталогов, обмен библиографическими базами данных, взаимное предоставление библиотеками доступа к своим информационным ресурсам, выдвинули на первый план новую функцию предметной рубрики – коммуникативную, роль которой заключается в обеспечении быстрого поиска нужного документа, то есть при однократной формальной обработке содержания каждого документа создается возможность многократного поиска информации о нем в базах данных. Сегодня вопросу унификации в предметизации уделяют особое внимание как одному из важнейших. Под унификацией понимают разработку общих и специальных принципов предметизации, создание авторитетных файлов предметных рубрик для единообразной обработки документа и минимизации субъективных действий каталогизаторов. Авторитетные файлы предметных рубрик обеспечивают возможность глубокого и результативного поиска в электронном каталоге, гарантируют его полноту и точность, служат контрольным словарем в организации тематического доступа к библиографическим записям.

Авторитетный/нормативный контроль предметных рубрик – это процесс поддержания единых форм авторитетных/нормативных предметных заголовков, которые означают один и тот же предмет в библиографическом файле, а также контроль за адекватностью присвоения предметного заголовка, за последовательным соблюдением принципов, методик, инструкций и правил для представления поисковых признаков документов.

Авторитетный/нормативный файл «Предметные заголовки» – это машиночитаемый словарь предметных заголовков, представленный в коммуникативном формате для авторитетных данных, с установленными взаимосвязями между предметными заголовками, выраженными элементами справочного аппарата, введенный в качестве базы знаний в подсистему лингвистического обеспечения электронного каталога и задействованный в процессе предметизации документов и поиска информации. Авторитетный файл «Предметные заголовки» является мощным инструментом не только для каталогизаторов в процессе предметизации документов, но и, что более важно, при организации тематического поиска в электронном каталоге.

В процессе предметизации наличие точных, конкретных, четких формулировок тер-

минов является очень важным фактором как для раскрытия содержания документов, так и для обеспечения максимальной эффективности тематического поиска, имеет большое значение, в частности, для эффективного кооперативного взаимодействия библиотек. В условиях быстрого накопления объемов информации в электронных каталогах и развития корпоративных межбиблиотечных объединений важным фактором успешного функционирования автоматизированной системы поиска и доставки документов является однообразие ввода информации. Решение этой проблемы мы видим во внедрении в процесс предметизации документов авторитетного контроля, который выполняет функцию организации и поддержки устойчивых форм заголовков и является эффективным инструментом унификации в предметизации. Обработка документов с использованием такого средства унификации, как авторитетные файлы предметных рубрик и их поддержание в актуальном состоянии гарантирует точность и полноту тематического поиска путем установления связей от принятых и утвержденных форм написания предметных заголовков к отклоненным формам заголовков. Это является обязательным условием успешного функционирования автоматизированной системы поиска документов, а также способствует интегрированию библиотек системы образования и развитию межбиблиотечного сотрудничества.

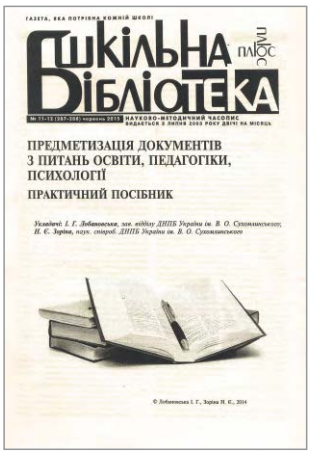
Не остались в стороне от этого процесса и специалисты ГНПБ Украины им. В.А. Сухомлинского. В течение 2009–2012 гг. при участии автора статьи были разработаны методические рекомендации «Предметизация документов по вопросам образования, педагогики и психологии», на основе которых в 2013 подготовлено практическое пособие с одноименным названием, опубликованное в 2014 г.

Неотъемлемой и очень важной составляющей пособия является Словарь предметных заголовков, который насчитывает около 2000 терминов и является базовым для формирования базы данных Авторитетного файла «Предметные заголовки» электронного каталога ГНПБУ. Отсутствие отечественных

В условиях быстрого накопления объемов информации в электронных каталогах и развития корпоративных межбиблиотечных объединений важным фактором успешного функционирования автоматизированной системы поиска и доставки документов является однообразие ввода информации

Таблица-1

ПРЕДМЕТНЫЙ ЗАГОЛОВОК	ТОЛКОВАНИЕ ТЕРМИНОВ	НЕПРИНЯТЫЕ ЗАГОЛОВКИ ПР, от которых формируются ссылки «См.» к принятому заголовку. Самостоятельные авторитетные записи не создаются	ОБОЗНАЧЕНИЕ	СВЯЗАННЫЕ ПРИНЯТЫЕ ЗАГОЛОВКИ ПР, к которым формируется ссылка «См. также» от принятого заголовка. Нужны самостоятельные авторитетные записи
А				
Академии	Разновидность научного или учебного заведения наивысшего уровня, а также некоторых обществ с почетным членством		Ш Ш Ш У У	Научные учреждения Высшие учебные заведения Учебные заведения Академии наук Духовные учебные заведения
Арт-педагогика	Направление педагогической науки, базирующееся на таких отраслях научного знания, как психология, искусство, педагогика. Функция А. – обеспечить разработку теории и практики личностно ориентированного эстетического образования, основная цель которого – развитие эмоционально-чувственной культуры и сохранения душевного покоя личности	Артпедагогика Естетопедагогика Психопедагогика искусства	Ш А А	Педагогика Педагогика искусства Естетическое воспитание
Арт-терапия	Направление и метод психотерапии, психокоррекции и реабилитации, особая форма терапии творчеством, связанная главным образом, с так называемыми визуальными искусствами – живописью, графикой, фотографией, скульптурой, а также их различными комбинациями с другими формами творческой деятельности	Арттерапия Артотерапия	Ш А А	Психотерапия Музикотерапия Сказкотерапия





В процессе формирования Словаря предметных заголовков авторы старались максимально учесть потребности как каталогизаторов, так и пользователей, обеспечить максимальную точность и научность формулировок предметных рубрик, удобство пользования и возможность эффективного поиска для пользователей разного уровня подготовленности

аналогов и насущная необходимость в нем обусловили актуальность этой работы. Это первая попытка создания отечественного отраслевого украиноязычного словаря для предметизации документов по вопросам образования, педагогики, психологии и смежных областей знаний, в котором собраны соответствующие термины, и на основе которого можно формировать авторитетные файлы предметных заголовков, осуществлять предметное индексирование документов и поиск нужной информации. В процессе формирования Словаря предметных заголовков авторы старались максимально учесть потребности как каталогизаторов, так и пользователей, обеспечить максимальную точность и научность формулировок предметных рубрик, удобство пользования и возможность эффективного поиска для пользователей разного уровня подготовленности. В Словаре предметных заголовков составители стремились максимально придерживаться требований, предъявляемых к языку предметных рубрик, поскольку именно использование авторитетного файла предметных заголовков поможет полно и точно передавать содержание документов в электронном каталоге; гарантировать однозначность толкования терминов индексирования, то есть иметь точную семантику, без синонимии, полисемии и омонимии; допускать многоаспектное индексирование; обеспечивать достаточную простоту и удобство индексирования, поиска информации и ведения электронного каталога; фиксировать парадигматические (логические и ассоциативные) связи, в том числе с помощью справочного аппарата; отображать современное состояние терминосистемы в области образования, педагогики, психологии и смежных дисциплин; давать возможность внесения исправлений и дополнений. В Словаре предметных заголовков использовали современную научную терминологию на украинском языке, за исключением понятий, более известных пользователям в иноязычном варианте, или тех, которые не имеют эквивалентов в украинском языке. Для определения формы унифицированных заголовков использовали авторитетные источники – официальные документы (государственные стандарты, законы Украины), словари, энциклопедии, справочники, научные публикации и др.

Список предметных заголовков в Словаре предметных заголовков представлен в виде таблицы, все составляющие которой используются в процессе создания авторитетной записи для каждого предметного заголовка:

- принятые предметные заголовки;
- определения терминов, пояснения к специальным терминам и словам иноязычного происхождения;
- непринятые заголовки – отклоненные термины (синонимы, другие формы написания слов, прямая или инверсированная форма словосочетаний и т.д.), от которых формируют ссылки «См.» к принятому заголовку;
- связанные принятые заголовки – к которым от принятого предметного заголовка делают ссылки «См. также» и каждый из которых требует создания самостоятельной авторитетной записи;
- условные обозначения, указывающие на иерархические и ассоциативные отношения между принятыми терминами (например, Ш – более широкое понятие, У – более узкое, А – ассоциативное) (таблица 1).

Каждый термин в таблице помещен в отдельной строке с указанием соответствующей информации в каждом столбце таблицы 1 (фрагменты Словаря предметных заголовков в переводе с украинского языка), например:

Ссылка «См.» – это ссылка от непринятых (отклоненных) терминов и формулировок к принятым в конкретной предметизационной системе. Ссылки «См.» несут основную терминологическую нагрузку, поскольку направляют поиск от возможных, но отклоненных формулировок предметных заголовков к единой принятой формулировке термина, соответствующей запросу. В Словаре предметных заголовков мы старались максимально учесть все возможные формы термина, под которым пользователь мог бы искать нужную тему: непринятые транскрипции конкретного термина, синонимы, сокращенные формы или аббревиатуры, инверсированные формы и т. Все непринятые формы термина отображают в Авторитетном файле «Предметные заголовки и с помощью ссылки «См.» направляют пользователя к нужной ему теме в принятой формулировке. Для непринятых форм предметного заголовка самостоятельные авторитетные записи не создаются.

Важной составляющей авторитетной записи являются связующие ссылки «См. также», указывающие на тематические связи между отдельными предметными рубриками, принятыми в конкретной предметизационной системе. С их помощью пользователь получает возможность расширить поиск за счет

других рубрик, с которыми связана предметная рубрика, по которой осуществляют поиск. Ссылки «См. также» отображают иерархические, ассоциативные, полииерархические и хронологические связи между принятыми заголовками и позволяют пользователю получить исчерпывающую информацию о наличии предметных рубрик, тематически близких к нужной ему предметной рубрике, и их связей между собой.

Такая структура авторитетной записи позволяет разместить практически исчерпывающую информацию по каждой предметной рубрике и предоставляет широкий диапазон возможностей для предметного поиска в электронном каталоге.

Для предотвращения неоправданного увеличения объема базы данных, повышения уровня унификации и уменьшения информационного шума нами было принято методическое решение о создании в базе данных «Авторитетный файл «Предметные заголовки» преимущественно одноуровневых записей без использования тематических, географических, хронологических и формальных подзаголовков. Подполя «Тематический подзаголовок», «Географический подзаголовок», «Хронологический подзаголовок» и «Формальный подзаголовок» каталогизаторы заполняют непосредственно во время обработки документов с использованием соответствующих словарей в процессе формирования предметных рубрик после ввода предметного заголовка через авторитетный файл.

В 2014 г. специалистами отдела комплектования и научной обработки документов ГНПБУ была начата работа по формированию в электронном каталоге библиотеки базы данных «Авторитетный файл «Предметные заголовки» с использованием Словаря предметных заголовков. Создание каждой авторитетной записи – это достаточно трудоемкий процесс, который требует не только высокого уровня внимания и сосредоточенности, но и научного подхода. Сегодня база данных «Авторитетный файл «Предметные заголовки» насчитывает около 1600 авторитетных записей предметных заголовков, большая часть которых используется в процессе предметизации документов в электронном каталоге. В настоящее время в связи с реформированием системы образования Украины и изменениями в украинском законодательстве в сфере образования и науки база данных «Авторитетный файл «Предметные заголовки» пополняется новыми терминами, а также осуществляется редактирование уже существующих авторитетных записей, уточнение терминов и их определений.

В перспективе, после полного завершения формирования базы данных «Авторитетный файл «Предметные заголовки», предметизацию документов в электронном каталоге ГНПБ Украины им. В.А. Сухомлинского планируется осуществлять исключительно с использованием авторитетного файла.

В заключение отметим, что эффективность использования информационных ресурсов библиотек и качество библиотечного обслуживания как на локальном уровне, так и в условиях межбиблиотечного взаимодействия в значительной мере зависит от качества используемых лингвистических средств тематического поиска в электронных каталогах. Применение такого средства унификации как Авторитетный файл «Предметные заголовки» для организации тематического поиска в электронном каталоге создаст условия для обеспечения максимальной точности, конкретности, четкости в процессе предметных документов, раскрытия их содержания, а также расширит возможности для достижения максимальной эффективности тематического поиска в электронных каталогах.

В 2014 г. специалистами отдела комплектования и научной обработки документов ГНПБУ была начата работа по формированию в электронном каталоге библиотеки базы данных «Авторитетный файл «Предметные заголовки» с использованием Словаря предметных заголовков

Сегодня база данных «Авторитетный файл «Предметные заголовки» насчитывает около 1600 авторитетных записей предметных заголовков, большая часть которых используется в процессе предметизации документов в электронном каталоге

Наталия Зорина, В.А. Сухомлинский номидаги Украина Давлат илмий педагогика кутубхонасининг Китоб тўкислаш ва ҳужжатларга илмий ишлов бериш бўлими илмий ходими

Муаллиф мақолада автоматлаштирилган ахборот-қидирув тизимларидан фойдаланишда электрон каталоглар ва библиографик маълумотлар базасини шакллантириш ва фойдаланувчилар эҳтиёжини қондириш мақсадида сифат жиҳатидан янги босқичда амалга оширилаётган ишлар ҳақида маълумот беради. Бугунги кунда ахборот олиш имкониятларига кўра, амалдаги анъанавий каталогларга нисбатан электрон каталогларнинг роли сезиларли даражада кенгайиб бормоқда. Бундай жараённи кескин ўзгаришида каталогизатор мутахассислар олдига янги вазифалар қўяди. Китобхон, энг аввало, электрон каталог орқали ахборотларни қидиришда минимал самара берувчи вақт сарфлаши зарур ва шу билан бир қаторда, фойдаланувчи сутка давомида интернет орқали ўзига керакли маълумотни хоҳлаган вақтида қидириш имкониятига эга бўлиш кераклиги ҳақидаги фикр-мулоҳазалари билан ўртоқлашган. ■