

СОЗДАНИЕ СОВРЕМЕННОЙ ИНТЕЛЛЕКТУАЛЬНОЙ ИНФОРМАЦИОННОЙ СИСТЕМЫ ДЛЯ ХРАНЕНИЯ, КАТАЛОГИЗИРОВАНИЯ И ПОИСКА НАУЧНОЙ ИНФОРМАЦИИ

Лысенко А.И.

*Лысенко Алексей Иванович – студент,
специальность: медицинская биофизика,
Медико-профилактический факультет,*

Новосибирский государственный медицинский университет, г. Новосибирск

***Аннотация:** на данный момент научные сообщества развитых стран мира обладают обширными информационными системами, которые могут в той или иной степени удовлетворять потребности пользователей в информации. Однако основными недостатками этих систем является не своевременная актуализация информации (не относится к библиотечным системам) и ограниченная возможность интеграции ресурсов как внутри системы, так и со сторонними системами (низкая интероперабельность).*

***Ключевые слова:** информационная система, интеллектуальный, научная статья, научный журнал, каталогизирование.*

Успешность деятельностью программиста определяется эффективностью работы компьютеров. То есть, прогресс первого напрямую зависит от того, удалось ли добиться, чтобы машины брали на себя часть интеллектуальной нагрузки. Максимальные результаты в современном мире позволяют достигнуть компьютеры, основанные на технологии искусственного интеллекта. Этот термин предполагает, что автоматизированные системы не только выполняют однотипные операции, но и обладают способностью к самообучению. Появление технологий искусственного интеллекта открывает перед всем человечеством множество направлений развития.

Одним из направлений в этой области является разработка интеллектуальных информационных систем. Последние представляют собой очередной эволюционный шаг в развитии традиционных систем. Они включают в себя наиболее наукоемкие технологии, посредством которых не только подготавливаются задачи по автоматизации процессов, но и автоматически разрабатываются варианты решений подобных целей. Причем системы определяют нужное направление, основываясь на ранее полученных данных.

Эта исследовательская работа призвана обосновать комплекс создания информационной системы, основанной на интеллектуальных технологиях и предназначенной для хранения, каталогизации и поиска научных данных.

Доступ к информационным технологиям и системам, а также их владение, повышает уровень конкурентноспособности специалиста на рынке труда в области автоматизации и реализации бизнес-процессов[1].

Степень развития информационных систем в современном мире определяется уровнем их интеллектуализации. Сегодня интеллектуальные технологии разрабатываются для достижения двух целей. Во-первых, они призваны расширить список задач, для решения которых привлекаются компьютеры. Достижение этой цели особенно актуально для слабоструктурированных предметных областей. Во-вторых, данные технологии способствуют повышению уровня интеллектуальной и информационной поддержки специалистов.

Рассматриваемые системы классифицируются на несколько типов в зависимости от степени их

реализации, сложности программного продукта, уровня интеграции с другими автоматизированными системами, вида компьютеров или сферы применения, особенностей решаемых задач, характера развития.

Квалификация человека в обществе определяется различными показателями. Но наивысшей степенью считается наличие способностей у специалиста генерировать знания (особенно новые). Достигнуть таких результатов позволяют интеллектуальные системы. К числу последних относятся технологии и средства, улучшающие качество хранения и обработки накопленной информации, систематизация знаний, решение комплексных и сложных задач, которые предусматривают одновременные обработку, накопление и пополнение данных.

В ходе планирования направлений развития высшего образования одним из ключевых моментов, позволяющих достигнуть нужные цели, считаются информационные технологии. Последние не только повышают уровень конкурентоспособности вузов на соответствующем рынке, но и обеспечивают результативную работу подобных учреждений. Без интеллектуальных технологий в современном обществе невозможно организовать эффективно функционирующий образовательный процесс [2].

Существует 4 важных основания, обуславливающих необходимость внедрения информационных технологий в образовательную сферу: профессиональные, социальные, педагогические и каталитические.

Для естественного вхождения высшей школы в мировое образовательное пространство нужно серьезно модернизировать и переориентировать компьютерно-информационную часть учебного процесса. Особенно важно найти решение вопросов, возникающих по мере внедрения автоматизированных процессов в обучение.

В настоящее время объем электронного документооборота растет по экспоненте. Этот факт демонстрирует несовершенство традиционных механизмов обработки документации: они не соответствуют современным потребностям конечных пользователей. Данная проблема затрагивает как глобальную, так и корпоративные сети.

Развитие технологий привело к возрастанию популярности информационных порталов. Последние создаются с целью консолидации информации и знаний.

К числу подобных решений относится исследовательский портал. Он представляет собой информационную систему, предназначенную для сбора и аналитической обработки поступающих данных, которые позволяют оценить инновационную активность регионов. Внедрение такого портала («Инновационное развитие регионов») увеличивает уровень поддержки, необходимой для принятия результативных управленческих решений. Данные, которые анализирует система, поступают из не- либо слабоструктурированных источников. В частности, к их числу относятся ресурсы глобальной Сети и оперативные базы данных.

Большинство технологий, предназначенных для обработки электронной документации, призваны повысить удобство работы с информацией. Однако нередко подобные решения являют собой аналог подходов, которые ранее применялись при обработке бумажных носителей.

Например, текстовые редакторы предоставляют множество возможностей для редактирования информации, благодаря чему последняя выстраивается в удобном для человеческого восприятия виде. Но подобные программы не способны передать смысловое содержание текста. То есть, в текстовых редакторах не реализовано семантическое индексирование.

Чтобы повысить эффективность процесса поиска решения данной проблемы, необходимо расширить понятие «документ». Последнее следует связывать со знаниями, позволяющими интерпретировать и обрабатывать данные, которые хранит этот документ[3].

Современные электронные документы преимущественно содержат в себе неструктурированную информацию. Задачей интеллектуальной системы, которая классифицирует и каталогизирует документы, заключается в следующем: она должна разбивать поступающую информацию по соответствующим темам

в рамках выбранных пользователей рубрик.

При этом процесс автоматической каталогизации регулярно сталкивается с рядом проблем (особенно актуально для документов в сети Интернет). Интеллектуальным системам приходится обрабатывать большой массив данных. Кроме того, внедрение автоматической каталогизации осложнено отсутствием структур, которые отслеживают появление в Сети новой информации, и авторской классификации. Под последним понимается приписывание к документам определенных кодов, аннотирование и выполнение иных действий, позволяющих быстро выявлять конкретный документ и массива данных. Еще одной сложностью, которая возникает в ходе внедрения автоматической каталогизации, является отсутствие возможности отслеживать изменения, которые вносятся в документы.

Существует два различных подхода к рассматриваемому процессу. Первый, считающийся наиболее эффективным (но тяжело реализуемым), предусматривает проведение каталогизации на основании знаний. Для этого используются ранее созданные базы знаний, в состав которых входит описание языковых выражений, соответствующие конкретной теме (рубрике), а также правила выбора самой рубрики[4].

Второй подход предполагает использование методов машинного обучения. В рамках данного решения компьютеры самостоятельно находят способы каталогизации на основании знаний, полученных в ходе изучения информации, которую вручную распределили по определенным рубрикам.

Внедряя систему автоматической каталогизации на ресурсе, необходимо решить следующие задачи:

1. Разработка механизма, отвечающего за введения новых рубрик и их описание. Последнее представляет собой некое выражение, основанное на словах и терминах, что встречаются в документах. Для решения этой задачи можно использовать экспертное описание рубрик либо методы машинного обучения, приведенные выше.

2. Анализ материала, учитывающий язык, на котором написан текст, и контекст употребления слов и фраз. Для решения этой задачи необходимо привлечь обширный перечень знаний, накопленных о языке и предметной области.

При разработке сложных программных продуктов длительного жизненного цикла создатели уделяют пристальное внимание вопросам квалифицированного сопровождения. В связи с этим возникает необходимость в реализации систем диалогового типа на основании интеллектуальных технологий. Такой подход позволяет обеспечить необходимый уровень сопровождения программного продукта[5].

Еще одним направлением, получившим широкое развитие, является создание и последующая интеграция баз технических данных в комплексные системы, которые разрабатываются с целью проектирования, производства технических объектов и контроля за их функционированием.

В настоящий момент разработчики проводят исследования, позволяющие решить следующие задачи:

- создание и последующая адаптация интеллектуальных компонентов диагностики;
- создание гибридных моделей экспертных систем;
- разработка новых технологий, позволяющих извлекать знания и определять закономерности, посредством которых можно группировать информацию;
- создание технологий, в рамках которых реализуются интеллектуальные системы имитационного моделирования;
- разработка методов, позволяющих проектировать и внедрять информационные системы и интеллектуальные человекомашинные интерфейсы;
- создание новых нейрокомпьютеров и нейроимитаторов с последующим внедрением в информационные системы[6].

В итоге получается, что интеллектуальные информационные системы представляют собой активно развивающуюся область информатики, которая имеет собственную функциональную основу и

настройку. Данное направление продолжит обновляться и расширяться с различной интенсивностью.

Приведенные ранее подходы можно применять при создании подсистем, ответственных за управление электронными документами на исследовательском портале. Важной особенностью описанных методов является то, что они ориентированы на использование онтологий в ходе представления знаний. Приведенные решения открывают дополнительные возможности для реализации интеллектуальных сервисов, предназначенных для поиска, обработки и каталогизации информации, получаемой из разных источников, в соответствии с тематикой ресурса.

Использование данных подходов позволяет решить следующие задачи:

- провести семантическое индексирование документов и внедрить интеллектуальный поиск информации, отвечающий потребностям конечных пользователей и специфике ресурса;
- извлечь нужных данных из неструктурированных документов;
- провести классификацию и каталогизацию данных в автоматическом режиме.

Список литературы

1. *Ланин В.В.* Интеллектуальное управление документами как основа технологии создания адаптируемых информационных систем // Труды международной научно-технической конференции «Интеллектуальные системы» (AIS'07). Т. 2 / М.: Физматлит, 2010. С. 334-339.
2. *Тарасов В.Б.* От многоагентных систем к интеллектуальным организациям: философия, психология, информатика. М.: Эдиториал. УРСС, 2013.
3. *Хорошевский В.Ф., Гаврилова Т.А.* Базы знаний интеллектуальных систем. СПб.: Питер, 2012.
4. *Ландэ Д.* Поиск знаний в Internet. Профессиональная работа. М.: Издательский дом «Вильямс», 2016.
5. *Ефремов В.* Search 2.0: огонь по «хвостам» // Открытые системы. СУБД №08 (134), 2007.
6. *Черняк Л.* Корпоративный поиск 2.0 // СУБД, 2014. № 07 (133).