

Департамент образования города Москвы
Государственное бюджетное образовательное учреждение
высшего профессионального образования города Москвы
«Московский городской педагогический университет»

Институт математики и информатики
Кафедра прикладной информатики

«Управление организационной информацией и знаниями»

Тема: «Проект № 3. Процессы поиска информации»



Выполнила:
студентка группы
ПИ 5
Гусихина
Е.С.

Москва
2014

Поиск необходимых данных

2

Поиск - процесс, в ходе которого в той или иной последовательности производится соотнесение отыскиваемого с каждым объектом, хранящимся в массиве. Цель любого поиска заключается в потребности, необходимости или желании находить различные виды информации, способствующие получению лицом, осуществляющим поиск, нужных ему сведений, знаний и т.д. для повышения собственного профессионального, культурного и любого иного уровня; создания новой информации и формирования новых знаний; принятия управленческих решений и т.п.

Информационный поиск

- Термин "информационный поиск" (англ. "information retrieval") ввёл американский математик К. Муэрс. Он заметил, что побудительной причиной такого поиска является информационная потребность, выраженная в форме информационного запроса. К объектам информационного поиска К. Муэрс отнес документы, сведения об их наличии и (или) местонахождении, фактографическую информацию.

Поисковые системы

4

- Системы, обеспечивающие реализацию подобного поиска информации, называются поисковыми системами (ПС). В традиционных технологиях ПС представляют картотеки и каталоги, адресные и иные справочники, указатели, энциклопедии, справочный аппарат к изданиям и другие материалы.

Информационно-поисковая система

5

- Информационно-поисковая система трактуется и как система, обеспечивающая поиск и отбор необходимых данных на основе информационно-поискового языка и соответствующих правил поиска, а база данных - как совокупность средств и методов описания, хранения и манипулирования данными, облегчающих сбор, накопление и обработку больших информационных массивов. Организация различных БД отличается видом объектов данных и отношений между ними.

ИПС делятся на: традиционные (ручные, механические, электромеханические) и автоматизированные (электронные).

6

- Автоматизированные ИПС (АИПС), используют компьютерные программно-технические средства и технологии и предназначаются для нахождения и выдачи пользователям информации по заданным критериям. Определяющими для понимания методов автоматизации поиска являются два следующих фактора:
 - 1) сравниваются не сами объекты, а описания - так называемые "поисковые образы";
 - 2) сам процесс является сложным (составным и не одноактным) и обычно реализуется последовательностью операций.

- Поскольку АИПС инструмент, используемый человеком при поиске (а не интеллектуальным автомат для поиска информации - готовых решений задач основной деятельности), эффективность её использования зависит от того, насколько хорошо человек знает природу операционных объектов и свойства инструмента, посредством которого он работает с этими объектами.
- Информационный поиск подразумевает использование определённых стратегий, методов, механизмов и средств. Поведение пользователя, осуществляющего управление процессом поиска, определяется не только информационной потребностью, но и инструментальным разнообразием системы - технологиями и средствами, предоставляемыми системой.

- Стратегия поиска - общий план (концепция, предпочтение, установка) поведения системы или пользователя для выражения и удовлетворения информационной потребности пользователя, обусловленный как характером цели и видом поиска, так и системными "стратегическими" решениями - архитектурой БД, методами и средствами поиска в конкретной АИПС. Выбор стратегии в общем случае является оптимизационной задачей. На практике в значительной степени он определяется искусством достижения компромисса между практическими потребностями и возможностями имеющихся средств.

- Метод поиска - совокупность моделей и алгоритмов реализации отдельных технологических этапов: построения поискового образа запроса (ПОЗ), отбора документов (сопоставление поисковых образов запросов и документов), расширения и реформулирования запроса, локализации и оценки выдачи.
- Поисковый образ запроса - записанный на ИПЯ текст, выражающий смысловое содержание информационного запроса и содержащий указания, необходимые для наиболее эффективного осуществления информационного поиска.

- По используемым поисковым технологиям ИС можно разбить на 4 категории:
- 1. Тематические каталоги;
- 2. Специализированные каталоги (онлайновые справочники);
- 3. Поисковые машины (полнотекстовый поиск);
- 4. Средства метапоиска.

- Тематические каталоги предусматривают обработку документов и отнесение их к одной из нескольких категорий, перечень которых заранее задан. Фактически это индексирование на основе классификации. Индексирование может проводиться автоматически или вручную с помощью специалистов, просматривающих популярные веб-узлы и составляющих краткое описание документов-резюме (ключевые слова, аннотация, реферат).

- Специализированные каталоги или справочники создаются по отдельным отраслям и темам, по новостям, по городам, по адресам электронной почты и т. п.
- Поисковые машины (самое развитое средство поиска в Интернете) реализуют технологию полнотекстового поиска. Индексируются тексты, расположенные на опрашиваемых серверах. Индекс может содержать информацию о нескольких миллионах документов.

- При использовании средств метапоиска запрос осуществляется одновременно несколькими поисковыми системами. Результат поиска объединяется в общий, упорядоченный по степени релевантности список. Каждая система обрабатывает только часть узлов сети, что позволяет расширить базу поиска. К подобному классу можно отнести и "персональные программы поиска", позволяющие формировать свои собственные инструменты метапоиска (например, автоматически опрашивать часто посещаемые узлы).

- ИПС фактически являются системами информационного обеспечения и представляют собой базы и банки данных. В качестве объекта в них выступает индивид, организация, отрасль, регион и т.п. Субъектом информационного обеспечения является специалист-информатик, любой потребитель информации.