



Тема 3. Информационно- поисковые системы

- Информатика – наука о сборе, переработке, хранении информации и доведении ее до потребителя.

Информационно-поисковые система

Комплекс программных, логических и лингвистических средств, предназначенных для поиска и отбора по заданным в цифровой форме признакам документов и данных, хранимых на машиночитаемых носителях обычно в виде баз данных.

База данных (БД)

Организованная совокупность блоков информационных элементов, представленных на машиночитаемых носителях, предназначенных и пригодных для оперативного решения пользовательских, служебных и других задач с использованием средств вычислительной техники. Ведение БД осуществляется с помощью СУБД

Система управления базами данных (СУБД)

СУБД – программный комплекс, служащий для целей создания, ведения БД, включающих:

- Создание и изменение структуры БД
- Ввод информации
- Корректировка информации
- Поиск информации
- Вывод результатов поиска

БД:

- Библиографические
- Фактографические
- Реферативные
- Полнотекстовые

- База данных состоит из записей, каждая из которых описывает один объект (документ).
- Запись состоит из полей.
- ПОЛЕ - определенная строка символов, идентифицированная меткой, содержащая одно или более подполей. Поле - часть записи, соответствующая одной области библиографического описания, одной предметной рубрике, одному шифру хранения и т.д.
- Повторяющееся поле – поле, содержащее часть записи, которая входит больше одного раза.

- **МЕТКА** - совокупность трех цифровых символов, используемая для идентификации поля.
- **ПОДПОЛЕ** - определенная единица информации в пределах поля.
- **ЭЛЕМЕНТ ДАННЫХ** - наименьшая единица библиографической записи, подлежащая точной идентификации. В переменном поле элемент данных идентифицируется идентификатором подполя, с помощью которого образуется подполе.
- **ИДЕНТИФИКАТОР ПОДПОЛЯ** - код, идентифицирующий отдельные подполя внутри переменного поля, Состоит из двух символов.
- **ИНДИКАТОР** - символ, цифровой или буквенный, связанный с переменным полем, представляющий дополнительную информацию о содержании поля

- Прямой поиск – поиск с использованием словарей (индексов, инвертированных файлов)
- Словарь (индекс, инвертированный файл) – файл содержащий термины из документов со ссылками на эти документы.
- Последовательный поиск – поиск производится последовательным перебором всех документов в БД.

Логические операторы

- И and *
 - А и В – выдаются документы, содержащие термин А и термин В одновременно
 - ИЛИ or +
 - А или В – выдаются документы, содержащие хотя бы один из терминов: А либо В, причем документы, содержащие оба термина одновременно выдаются один раз.
 - Не not ^
 - А не В - выдаются документы, содержащие термин А, но не содержащие термин В.
 - \$ *
- Оператор правостороннего усечения: А\$ - выдаются документы начинающиеся с А, например: Иван\$ будут выданы документы с автором Иванов, Иванова, Иваненко и т. д.

- Благодарю за внимание!
- к. т. н. С. Р. Баженов, зав. отделом
автоматизированных систем ГПНТБ СО
РАН