

Информационный ПОИСК

Лекция 3

План

- Информационный поиск (ИП).
- Этапы ИП.
- Задачи ИП.
- Виды ИП.
- Информационно-поисковая система (ИПС).
- Запрос, объект запроса.
- Оценка эффективности ИП.

Поиск

совокупность операций, связанных с определением местонахождения объекта с заданными характеристиками или признаками

Поиск информации

- процесс выявления в некотором множестве документов тех, которые

✓ посвящены указанной теме

✓ удовлетворяют условию поиска

✓ содержат необходимые факты, сведения, данные

Информационный поиск

Кельвин Муэрс, 1948

процесс поиска (нахождения, отбора и выдачи) информации из массивов и записей любого вида и на любых носителях, которая определяется заранее заданными признаками и удовлетворяет информационные потребности

наука об этом
поиске

Причина осуществления ИП - **информационная потребность** в форме **информационного запроса**

Этапы ИП

уточнение потребности, формулировка запроса



определение источников



извлечение информации



ознакомление с полученной информацией,
оценка результатов

Цель ИП

- ПОМОЧЬ ПОЛЬЗОВАТЕЛЮ
удовлетворить его
информационную потребность.

Основные задачи ИП

1. Поиск необходимых сведений об источнике и установление его наличия в системе других источников (**библиографический поиск**).
2. Поиск информационных источников (документов и изданий), в которых содержится нужная информация (**документальный поиск**).
3. Поиск фактических сведений, содержащихся в литературе, книге (**фактографический поиск**).

Виды ИП (критерии выделения)

- Цель поиска
- Объект поиска
- Предмет поиска
- Степень использования технических средств
- Метод поиска
- Жанр литературы
- Хронологический охват
- Географический охват
- Полнота поиска
- Интенсивность поиска
- Категория потребителя информации
- Интенсивность поиска

Цель ИП

Адресный

- процесс поиска документов по формальным признакам, указанным в запросе

Условия:

- наличие у документа точного адреса
- строгий порядок расположения документов

Семантический

- процесс поиска документов по их содержанию.

Условия:

- перевод содержания с естественного языка на информационно-поисковый язык,
- составление поисковых образов
- составление поискового описания

Объект ИП

Документальный ИП

- **Библиотечный** -
нахождение первичных документов.
- **Библиографический** -
нахождение сведений о документах,
представленных в виде
библиографических
записей.

Фактографический ИП

- **Документально-фактографический** - поиск в документах фрагментов текста, содержащих факты.
- **Фактологический** - создание новых фактографических описаний путем переработки найденной фактографической информации.

Предмет (содержание) ИП

- Универсальный
- Отраслевой
- Тематический

Степень использования технических средств

- Ручной
- Машинный
- Автоматизированный

Методы ИП

- Сплошной поиск
- Выборочный поиск
- Интуитивный поиск
- Индуктивный поиск
- Дедуктивный поиск
- Рецептурный поиск
- Поиск по библиографическим ссылкам
- Поиск путем восхождения от абстрактного к конкретному

Жанр литературы

- Общелитературный
- Жанровый (видовой)

Хронологический охват

- Текущий
- Ретроспективный
- Проспективный (перспективный)

Географический охват

- Общемировой
- Страноведческий
- Краеведческий

Полнота поиска

- Абсолютный (максимальный)
- Избирательный

ИНТЕНСИВНОСТЬ поиска

- Систематический
- Разовый

Категория потребителя информации

- Массовый
- Специальный
- Индивидуальный

ИНТЕНСИВНОСТЬ поиска

- Разовый
- Многократный
- Постоянный

Информационно-поисковая система (ИПС)

автоматизированная поисковая система, реализованная на средствах электронной вычислительной техники и предназначенная для нахождения и выдачи пользователям необходимой информации по заданным критериям

ИПС=ИПЯ + программные средства + средства/правила перевода текстов на ИПЯ (индексирование)

Виды ИПС

- документографические (документальные)
- фактографические
- документально-фактографические (интегрированные)

Оценка эффективности ИТ

Критерий соответствия –
признак/совокупность признаков,
определяющих степень соответствия
между запросом и документом для
принятия решения о выдаче или не
выдаче документа в ответ на запрос

Критерий смыслового соответствия

признак/совокупность признаков, по которым определяется степень **смыслового** соответствия между **содержанием** запроса и документом для принятия решения о выдаче или не выдаче документа в ответ на запрос

Критерий формального соответствия

признак/совокупность признаков, по которым определяется степень **формального** соответствия между запросом и документом для принятия решения о выдаче или не выдаче документа в ответ на запрос

Релевантность

Коэффициент
потерь

Точность

Коэффициент
информационного
шума

Полнота

Релевантность

- характеристика степени соответствия смыслового содержания документа, найденного в результате поиска, содержанию информационного запроса.

Точность (precision)

- определяется путем деления количества выданных в результате выполнения поиска релевантных документов на общее число выданных документов (релевантных и не релевантных)

Полнота (recall)

- определяется путем деления количества выданных в результате выполнения поиска релевантных документов на общее число релевантных документов, имеющих в информационно-поисковой системе (выданных и не выданных)

Коэффициент информационного шума

- определяется путем деления количества выданных в результате выполнения поиска нерелевантных документов на общее число выданных документов (релевантных и не релевантных)

Коэффициент потерь

- определяется путем деления количества не выданных в результате выполнения поиска релевантных документов на общее число релевантных документов, имеющих в информационно-поисковой системе (выданных и не выданных)

Задание

Общее число релевантных документов в базе данных - 160.

Общее число найденных по запросу документов – 100.

Из них релевантных 40, нерелевантных 60.

Определите

- **ТОЧНОСТЬ**
- **ПОЛНОТУ**
- **КП**
- **КИШ**

Семинар 3

1. Информационный поиск (ИП): понятие ИП, ИП как наука и процесс
2. Этапы ИП
3. Задачи ИП
4. Виды ИП
5. Методы ИП
6. Информационно-поисковая система (ИПС). Виды
7. Оценка эффективности ИП
8. Поисковые образы документов

Литература

- Зубов А.В. Информационные технологии в лингвистике. Москва, 2004.
- Захаров В.В. Лингвистические средства информационного поиска в Интернете. Санкт-Петербург, 2005.
- Соснина Е.П. Введение в прикладную лингвистику. Ульяновск, 2012.