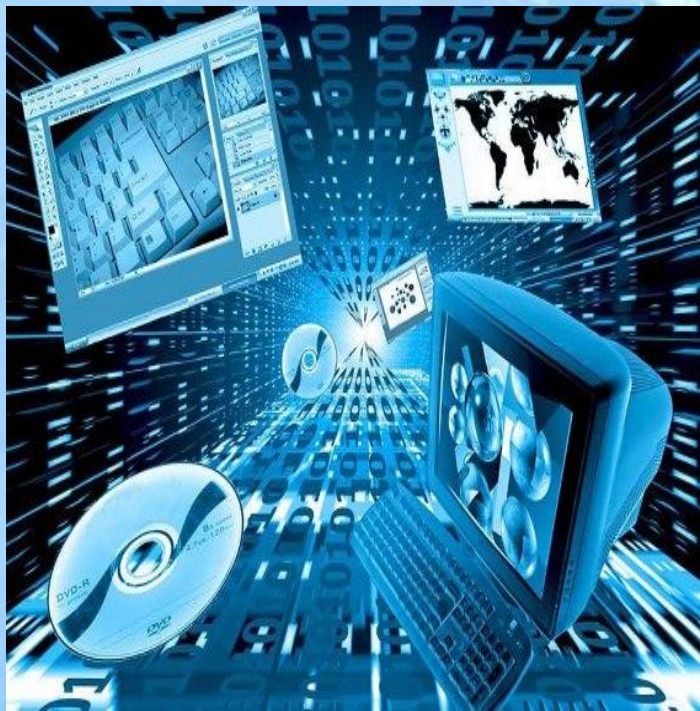


СВЯЗЬ ЛИНГВИСТИКИ И ИНФОРМАТИКИ



ЧТО ТАКОЕ ЛИНГВИСТИКА?



Лингвистика
Наука о языке

Лингвистика (от лат. *lingua* — язык), **языкознание**, **языковедение** — наука, изучающая языки.

Это наука о естественном человеческом языке вообще и обо всех языках мира как его индивидуализированных представителях.

В широком смысле слова, лингвистика подразделяется на научную и практическую. Чаще всего под лингвистикой подразумевается именно научная лингвистика. Является частью семиотики как науки о знаках.

Лингвистикой профессионально занимаются учёные-лингвисты.

Что такое информатика?

- Информатика – это техническая наука, определяющая сферу деятельности, связанную с процессами хранения, преобразования и передачи информации с помощью компьютера.
- Компьютер – универсальный прибор для обработки информации.



Связь правовой информатики с другими науками

Лингвистика – наука о языке

Методы лингвистики используются в информатике при разработке принципов автоматического перевода и информационно-поисковых языков, индексирования и реферирования, при составлении тезаурусов, упорядочении терминологии.



Лингвистика и Информатика.

В жизни современного общества важную роль играют автоматизированные информационные технологии. Но развитие информационных технологий происходит весьма неравномерно: если современный уровень вычислительной техники и средств связи поражает воображение, то в области смысловой обработки информации успехи значительно скромнее. Эти успехи зависят, прежде всего, от достижений в изучении процессов человеческого мышления, процессов речевого общения между людьми и от умения моделировать эти процессы на ЭВМ. А это - задача чрезвычайной сложности. Когда речь идет о создании перспективных информационных технологий, то проблемы автоматической обработки текстовой информации, представленной на естественных языках, выступают на передний план. Это определяется тем, что мышление человека тесно связано с его языком. Более того, естественный язык является инструментом мышления. Он является также универсальным средством общения между людьми – средством восприятия, накопления, хранения, обработки и передачи информации.

Проблемами использования естественного языка в системах автоматической обработки информации занимается наука компьютерная лингвистика. Эта наука возникла сравнительно недавно – на рубеже пятидесятих и шестидесятих годов прошлого столетия. Поначалу, в период своего становления, она имела различные названия: математическая лингвистика, вычислительная лингвистика, инженерная лингвистика. Но в начале восьмидесятих годов за ней закрепилось название компьютерная лингвистика.

Компьютерная лингвистика - это область знаний, связанная с решением задач автоматической обработки информации, представленной на естественном языке. Центральными научными проблемами компьютерной лингвистики являются проблема моделирования процесса понимания смысла текстов (перехода от текста к формализованному представлению его смысла) и проблема синтеза речи (перехода от формализованного представления смысла к текстам на естественном языке). Эти проблемы возникают при решении ряда прикладных задач:

- 1) автоматического обнаружения и исправления ошибок при вводе текстов в ЭВМ,
- 2) автоматического анализа и синтеза устной речи,
- 3) автоматического перевода текстов с одних языков на другие,
- 4) общения с ЭВМ на естественном языке,
- 5) автоматической классификации и индексирования текстовых документов, их автоматического реферирования, поиска документов в полнотекстовых базах данных.

За прошедшие полвека в области компьютерной лингвистики были получены значительные научные и практические результаты: были созданы системы машинного перевода текстов с одних естественных языков на другие, системы автоматизированного поиска информации в текстах, системы автоматического анализа и синтеза устной речи и многие другие. Но были и разочарования. Например, проблема машинного перевода текстов с одних языков на другие оказалась значительно сложнее, чем это представляли себе пионеры машинного перевода и их последователи. То же самое можно сказать об автоматизированном поиске информации в текстах и о задаче анализа и синтеза устной речи. Ученым и инженерам придется по-видимому еще немало потрудиться, чтобы достигнуть нужных результатов.

Обработка естественного языка ([англ. natural language processing](#); синтаксический, морфологический, семантический анализы текста). Сюда включают также:

Корпусная лингвистика, создание и использование электронных корпусов текстов
Создание электронных словарей, тезаурусов, онтологий. Например, [Lingvo](#). Словари используют, например, для автоматического перевода, проверки орфографии.

Автоматический перевод текстов. Среди русских переводчиков популярным является [Промт](#). Среди бесплатных известен переводчик [Google Translate](#)

Автоматическое извлечение фактов из текста (извлечение информации) ([англ. fact extraction, text mining](#))

Автореферирование ([англ. automatic text summarization](#)). Эта функция включена, например, в [Microsoft Word](#).

Построение систем управления знаниями. См. [Экспертные системы](#)

Создание вопросно-ответных систем ([англ. question answering systems](#)).

Оптическое распознавание символов ([англ. OCR](#)). Например, программа [FineReader](#)

Автоматическое распознавание речи ([англ. ASR](#)). Есть платное и бесплатное ПО

Автоматический синтез речи


```
#include <iostream>
#include <string>
int mainn()
{
    std::string name;
    std::cout << "Как тебя зовут? ";
    getline (std::cin, name);
    std::cout << "Привет, " << name << "!\n";
}
```

3:2: error: invalid preprocessing directive #include In function 'int mainn()':

11:1: warning: no return statement in function returning non-void [-Wreturn-type]