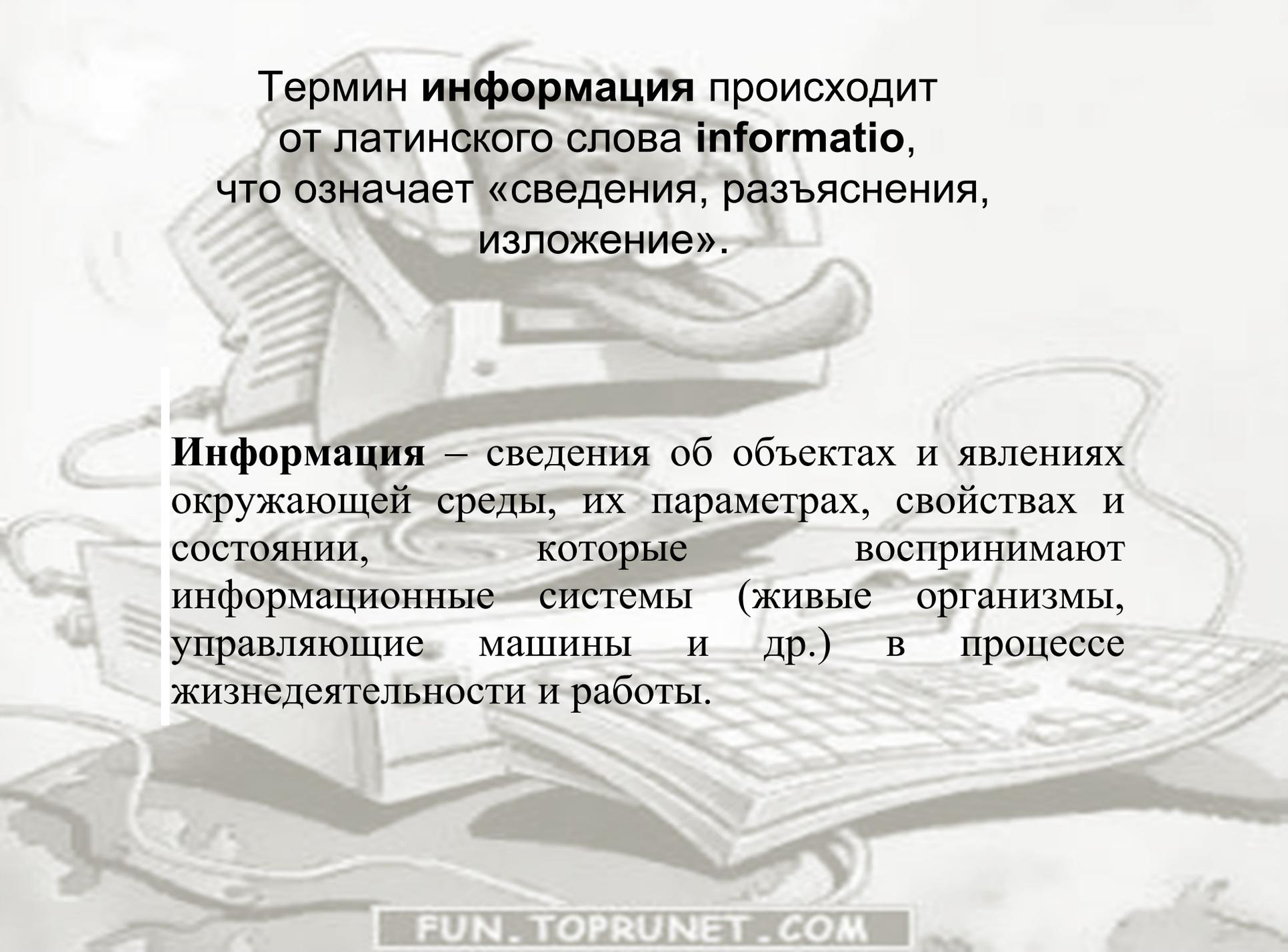


Учебная дисциплина

**Информационные технологии в  
лингвистике**

Информация, ее виды и свойства

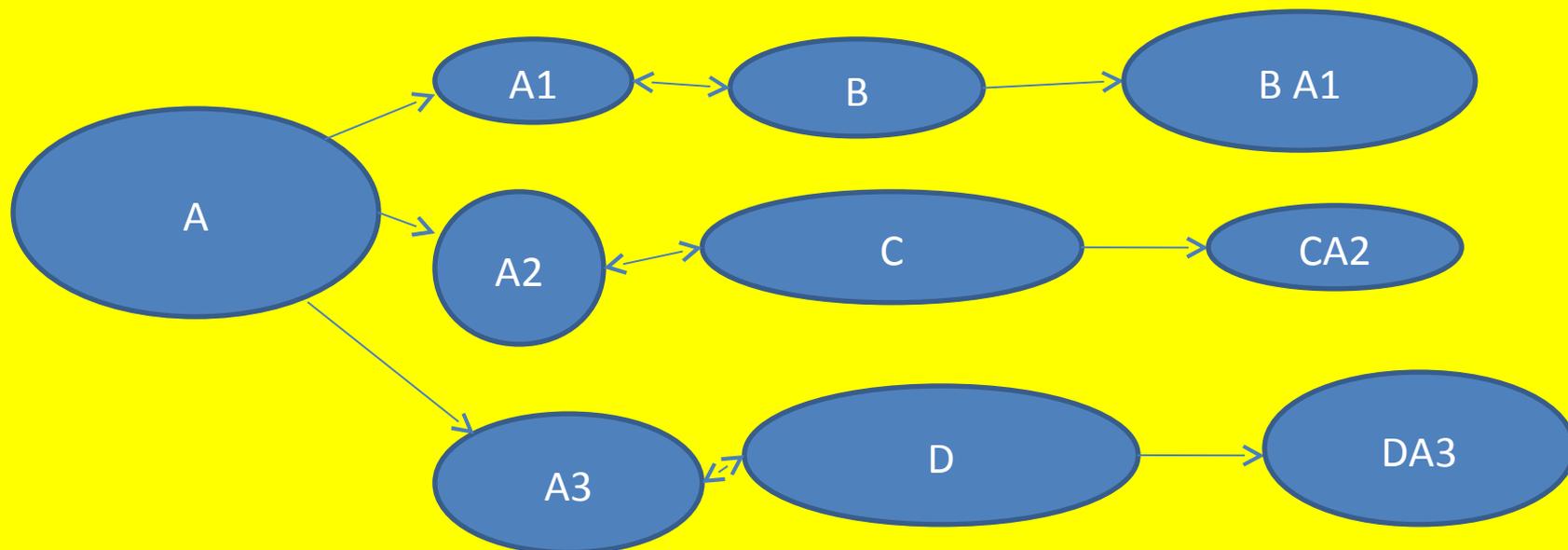


Термин **информация** происходит от латинского слова **informatio**, что означает «сведения, разъяснения, изложение».

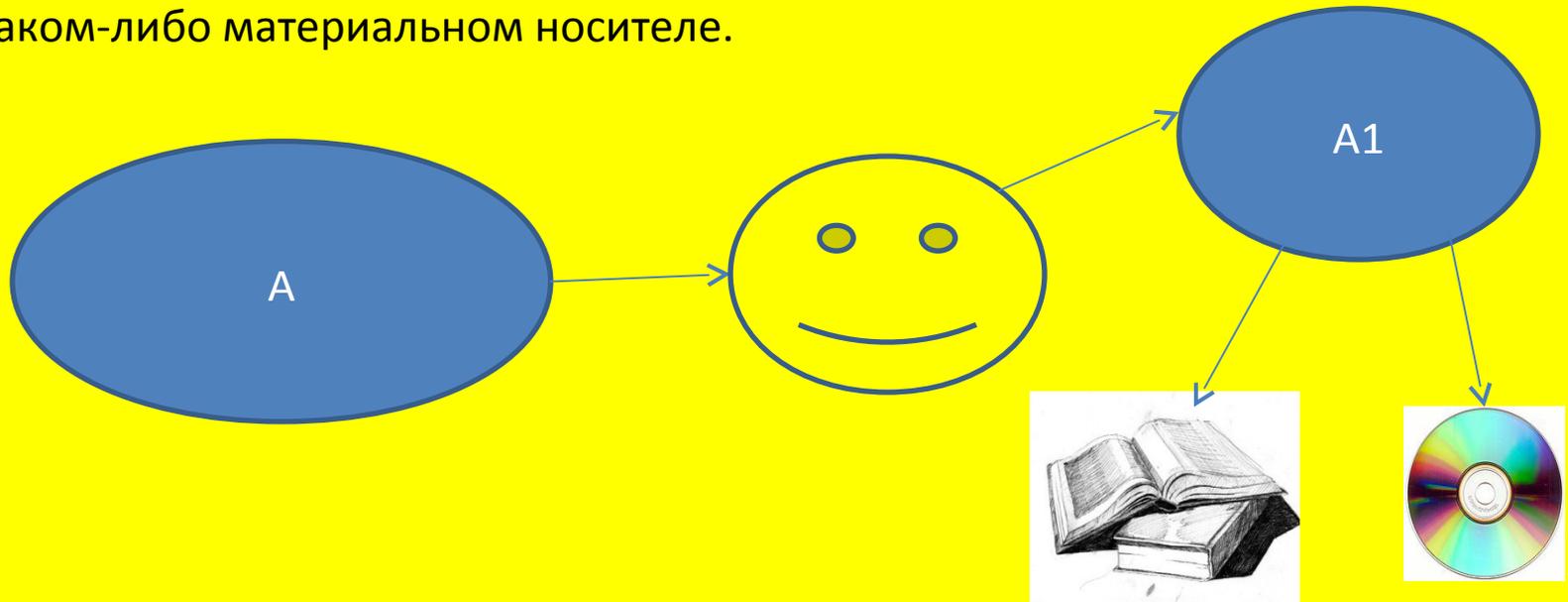
**Информация** – сведения об объектах и явлениях окружающей среды, их параметрах, свойствах и состоянии, которые воспринимают информационные системы (живые организмы, управляющие машины и др.) в процессе жизнедеятельности и работы.

**В современной науке рассматриваются два вида информации**

*Объективная (первичная) информация — свойство материальных объектов и явлений (процессов) порождать многообразие состояний, которые посредством взаимодействий (фундаментальные взаимодействия) передаются другим объектам и запечатляются в их структуре.*



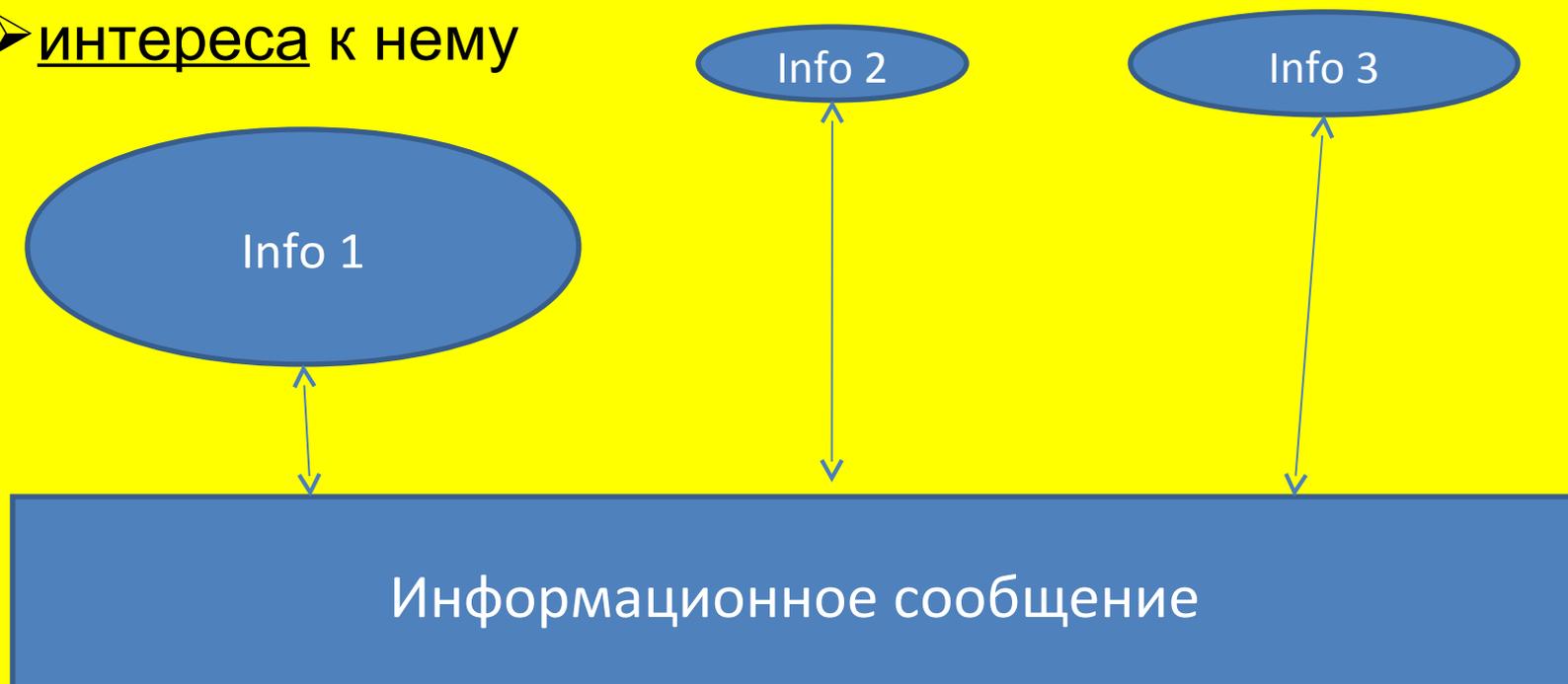
Субъективная (*семантическая, смысловая, вторичная*) информация – смысловое содержание объективной информации об объектах и процессах материального мира, сформированное сознанием человека с помощью смысловых образов (слов, образов и ощущений) и зафиксированное на каком-либо материальном носителе.



В бытовом смысле **информация** – сведения об окружающем мире и протекающих в нем процессах, воспринимаемые человеком или специальным устройством.

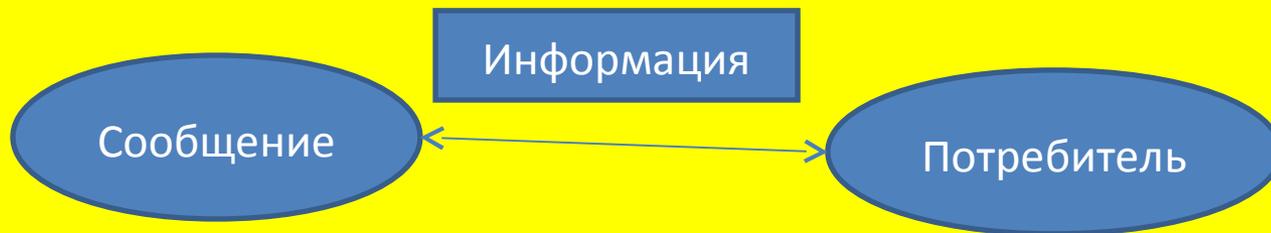
Одно и то же информационное сообщение (статья в газете, объявление, письмо, телеграмма, справка, рассказ, чертеж, радиопередача и т. п.) может содержать разное количество информации для разных людей в зависимости от

- их накопленных знаний,
- уровня понимания этого сообщения
- интереса к нему

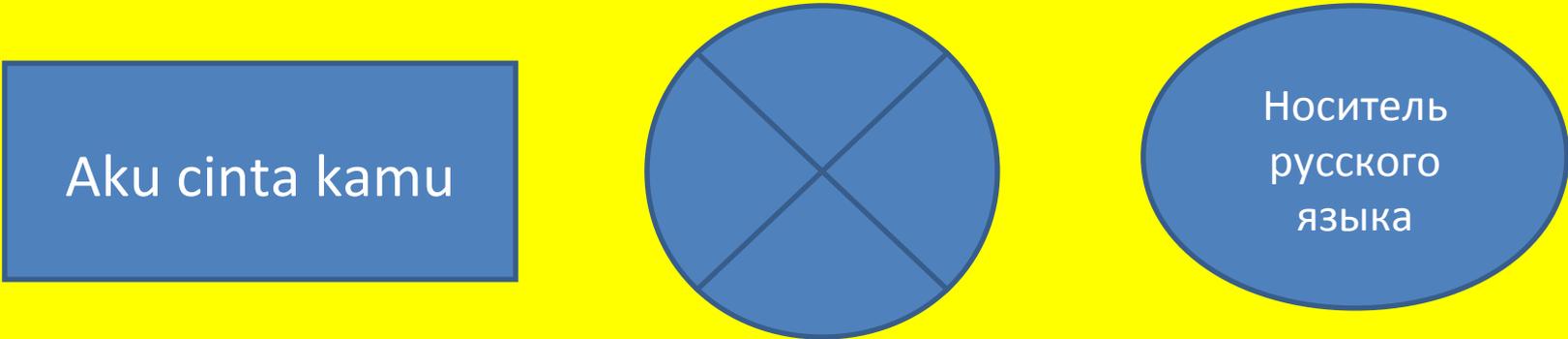


**Информация – характеристика не сообщения, а соотношения между **сообщением** и его **потребителем**.**

**Без наличия потребителя, хотя бы потенциального, говорить об информации бессмысленно.**



Сообщение, составленное *на определенном языке*, не несет никакой новой информации человеку, не знающему этого языка, но может быть высокоинформативным для человека, владеющего данным языком

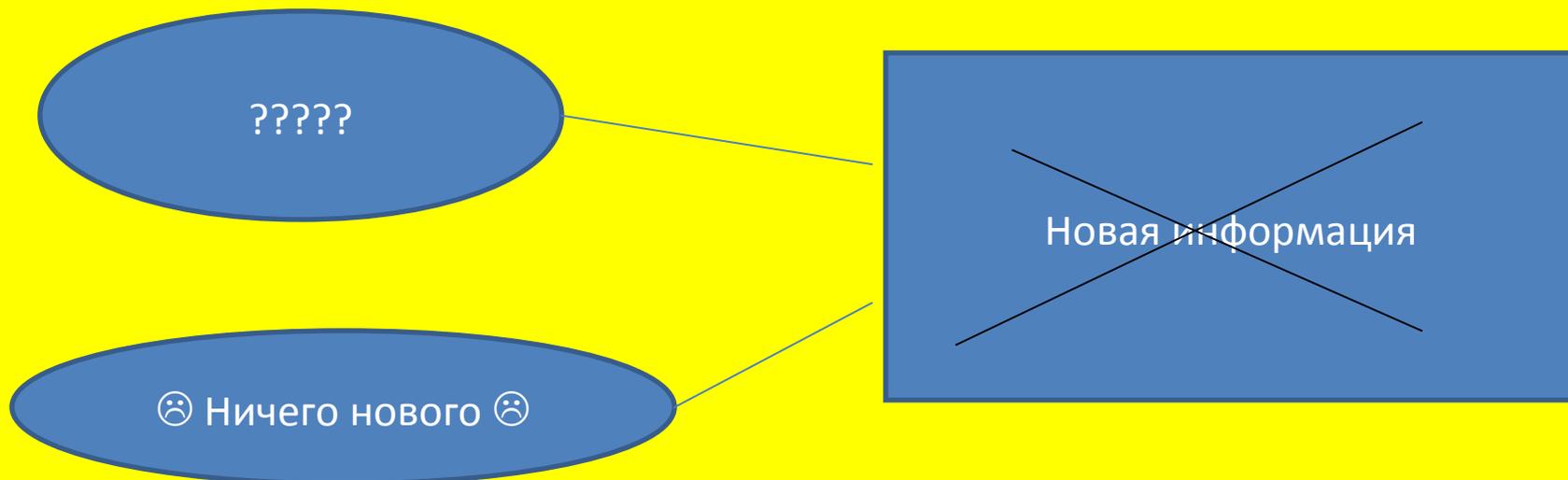


Aku cinta kamu

Носитель  
русского  
языка

*Никакой новой информации не содержит и сообщение, изложенное на знакомом языке, если его содержание*

*непонятно  
или  
уже известно*



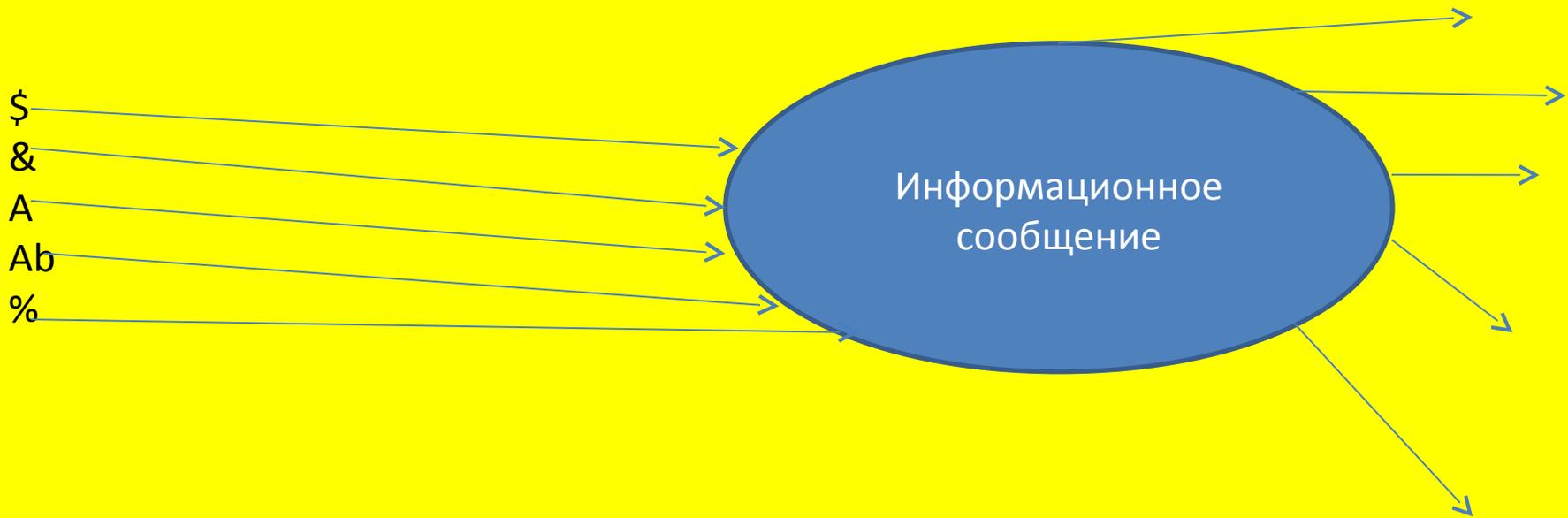
*Применительно к компьютерной обработке данных*

под **информацией** понимают

некоторую **последовательность символических обозначений** (букв, цифр, закодированных графических образов и звуков и т. п.), несущую смысловую нагрузку и представленную в понятном компьютеру виде



Каждый новый символ в такой последовательности символов увеличивает информационный объем сообщения



Информация может существовать в виде:

- текстов, рисунков, чертежей, фотографий;
- световых или звуковых сигналов;
- электрических и нервных импульсов;
- магнитных записей;
- жестов и мимики;
- запахов и вкусовых ощущений;
- хромосом, посредством которых передаются по наследству признаки и свойства организмов;
- радиоволн и т.д.

# Что можно делать с информацией???

создавать	принимать	комбинировать	хранить
передавать	копировать	обрабатывать	искать
воспринимать	формализовать	делить на части	измерять
использовать	распространять	упрощать	разрушать
запоминать	преобразовывать	собирать	и т. д.

➤ Достоверность



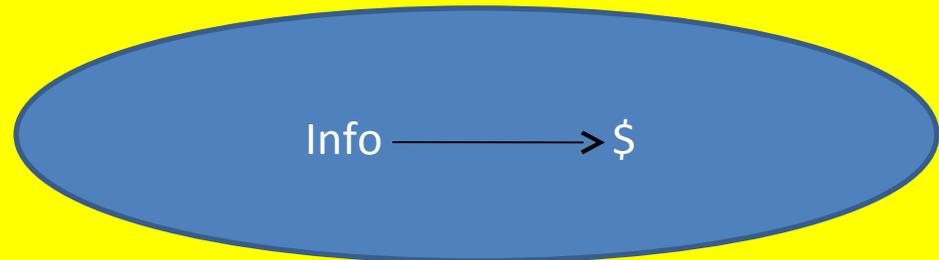
➤ Полнота



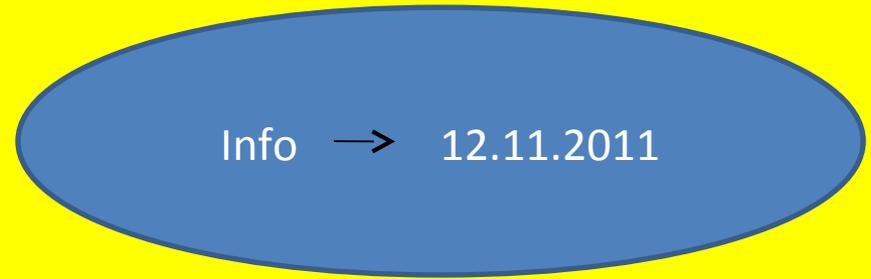
➤ Точность



➤ Ценность



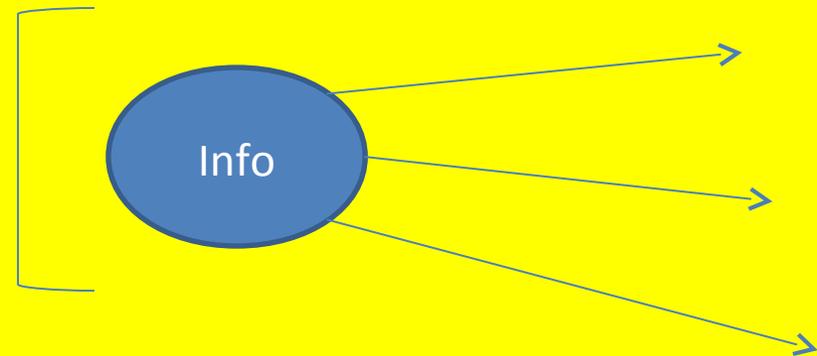
➤ Своевременность



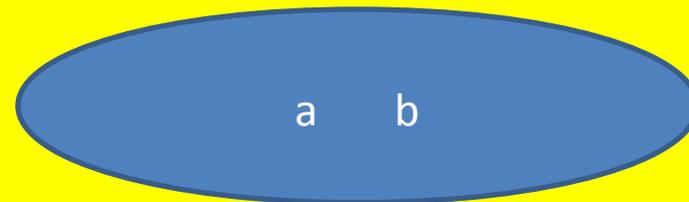
➤ Понятность



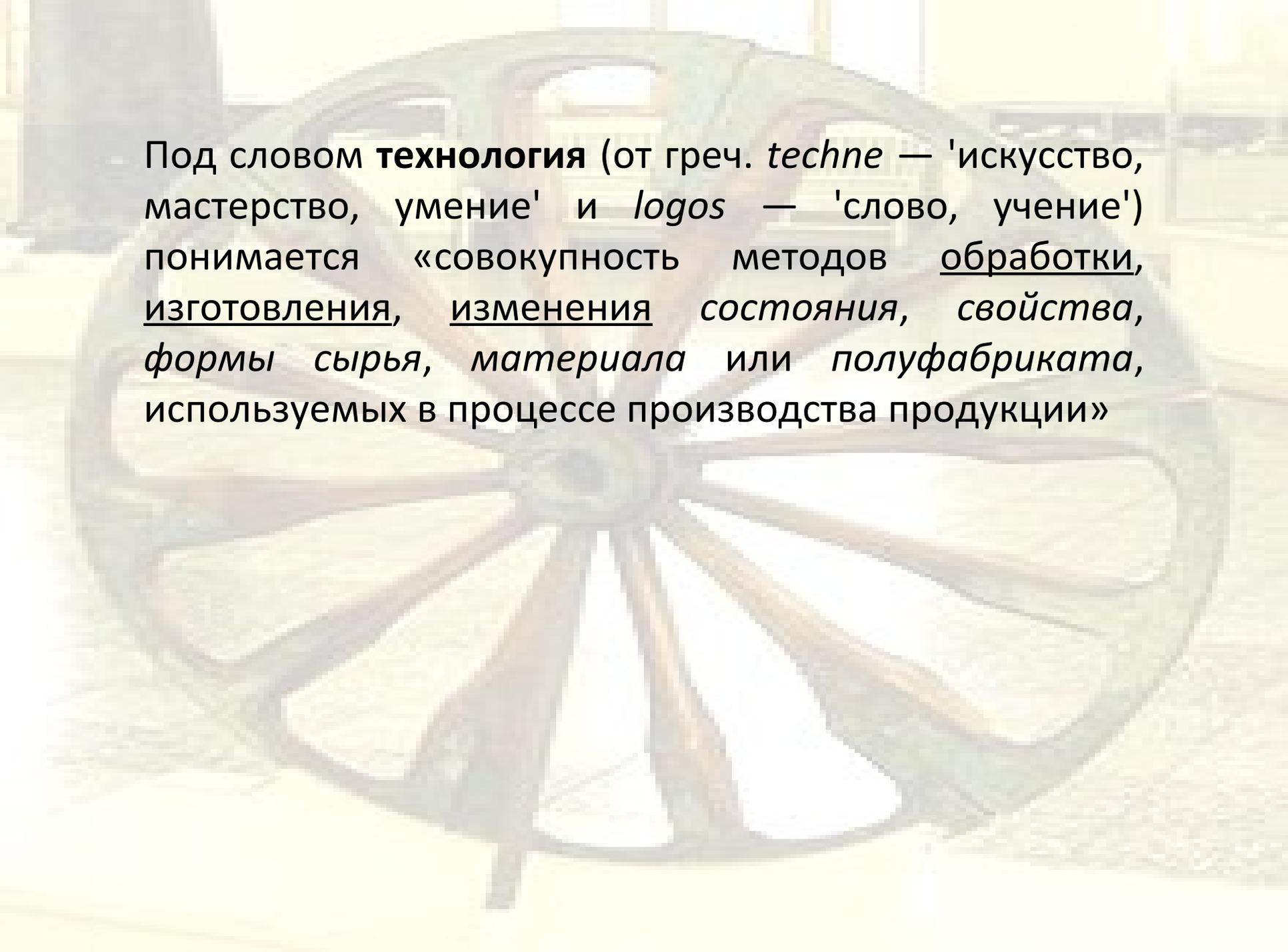
➤ Доступность



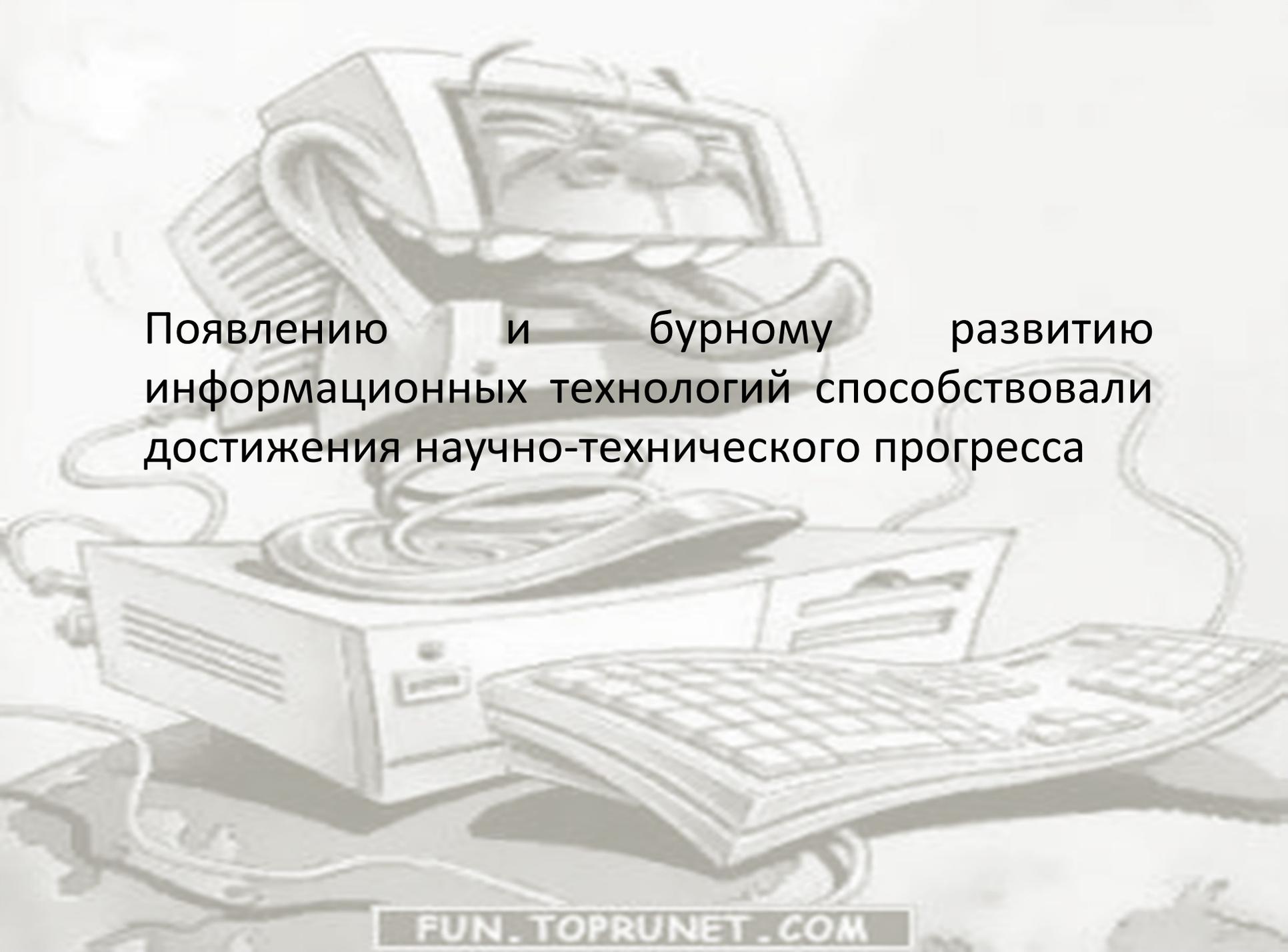
➤ Краткость и т. д.



**ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ  
И ПРИЧИНЫ,  
СПОСОБСТВУЮЩИЕ ИХ ПОЯВЛЕНИЮ**

A large, dark wooden water wheel is the central focus of the image. It has a circular frame with several thick wooden spokes radiating from a central hub to the outer rim. The wheel is positioned in front of a light-colored wall with a window. The floor is made of light-colored stone tiles. The overall scene is brightly lit, suggesting an indoor museum or historical site.

Под словом **технология** (от греч. *techne* — 'искусство, мастерство, умение' и *logos* — 'слово, учение') понимается «совокупность методов обработки, изготовления, изменения *состояния*, *свойства*, *формы сырья*, *материала* или *полуфабриката*, используемых в процессе производства продукции»



Появлению и бурному развитию информационных технологий способствовали достижения научно-технического прогресса

Создание персональных компьютеров с большой памятью и большой скоростью выполнения операций



Разработка звуковых плат, дающих возможность воспроизводить и записывать речь, звуки и музыку в большом диапазоне частот



Изобретение видеоплат, позволяющих выводить на экран компьютеров изображение с телеэкранов и видеомагнитофонов



Разработка специальных (мультимедийных) компьютеров, позволяющих воспроизводить на экране дисплеев цвет, звук, музыку, движение

Основные **составляющие** мультимедиа:



Текст



Изображение



Анимация



Видео

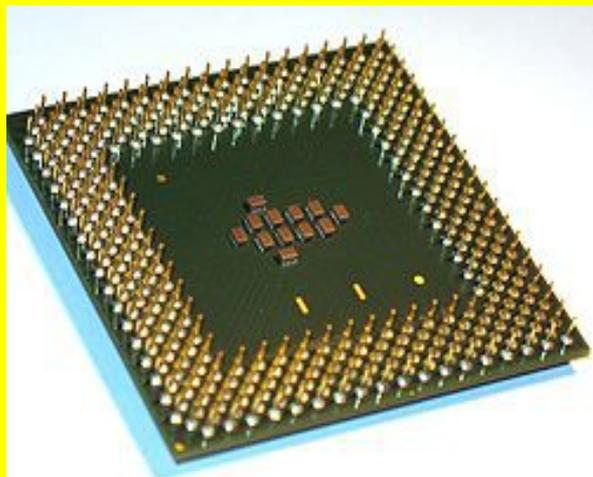


Аудио



Интерактивность

Создание специальных процессоров и устройств, способных передавать информацию в сети от одного компьютера к другому



Микропроцессор



Сетевой концентратор



Создание электронной оргтехники, связанной с персональными компьютерами и позволяющей осуществлять высокоскоростную печать документов, их копирование и размножение



# **ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ЛИНГВИСТИКЕ**

**Информационные технологии в лингвистике** — это совокупность *законов, методов и средств получения, хранения, передачи, распространения, преобразования* информации о языке и законах его функционирования с помощью компьютеров.

Понятие «информационные технологии» в лингвистике относится в основном к задачам прикладной лингвистики. К их числу можно отнести:

- создание систем искусственного интеллекта;
- создание систем автоматического перевода;
- создание систем автоматического аннотирования и реферирования текстов;
- создание систем порождения текстов;
- создание систем обучения языку;
- создание систем понимания устной речи;

- создание систем генерации речи;
- создание автоматизированных информационно-поисковых систем;
- создание систем атрибуции и дешифровки анонимных и псевдоанонимных текстов;
- разработка различных баз данных (словарей, карточек, каталогов, реестров и т.п.) для гуманитарных наук;
- разработка различного типа автоматических словарей;
- разработка систем передачи информации в сети Интернет и т.д.

Данные задачи включают целый ряд более мелких проблем. К их числу относится автоматизация следующих процессов:

- построение словарей текстов;
- морфологический анализ слова;
- определение значения многозначного слова;
- синтаксический анализ предложения;
- поиск слова в словаре;
- порождение предложения и т.д.

**Прикладная лингвистика** (*прикладное языкознание*) — наряду с теоретической лингвистикой является частью науки, занимающейся языком. Специализируется на решении практических задач, связанных с изучением языка, а также на практическом использовании лингвистической теории в других областях.

Основные направления прикладной лингвистики, связанные с изучением языка:

**лексикография** — теория и практика составления словарей;

**лингводидактика** — наука о разработках методик обучения иностранному языку;

**терминоведение** — наука об упорядочении и стандартизации научно-технической терминологии;

**переводоведение** — теория перевода.

Основные направления прикладной лингвистики (англ. *applied linguistics*), связанные с практическими приложениями:

Компьютерная лингвистика (англ. *computational linguistics*):

- Машинный перевод
- Автоматическое распознавание символов (англ. *OCR*)
- Автоматическое распознавание речи (англ. *ASR*)
- Автоматическое извлечение данных (англ. *Data Mining*)
- Автоматическое реферирование текстов
- Построение систем управления знаниями
- Создание электронных словарей, тезаурусов, онтологий
- Корпусная лингвистика, создание и использование электронных корпусов текстов

Лингвистическая экспертиза (например, в судебной практике)

наука об упорядочении и стандартизации научно-технической терминологии

Сайт международной конференции по  
компьютерной лингвистике

<http://www.dialog-21.ru/>