

# Математическая лингвистика

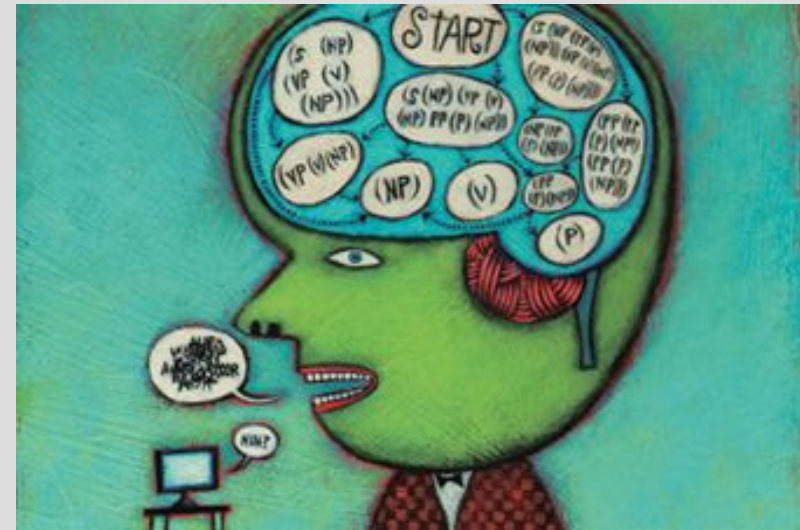


# Что такое математическая (компьютерная) лингвистика?

Это математическая дисциплина, предметом которой является разработка формального аппарата для описания строения естественных и некоторых искусственных языков.

## Задачи компьютерной лингвистики:

1. Распознавание и синтез речи
2. Машинный перевод
3. Анализ текста
4. Автоматическое извлечение данных



# Томи́та - парсер

Томи́та-парсер — это инструмент для извлечения структурированных данных (фактов) из текста на естественном языке. Извлечение фактов происходит при помощи контекстно-свободных грамматик и словарей ключевых слов.



# Как работает томика-парсер?

```
#encoding "utf-8"  
#GRAMMAR_ROOT S  
  
S ->Word<h-reg1, ~fw,wfm="[А-Я] [а-я, ё] *">;
```

«Князь Андрей, говоря это, был еще менее похож, чем прежде, на того Болконского, который развалившись сидел в креслах Анны Павловны и сквозь зубы, щурясь, говорил французские фразы. Его сухое лицо все дрожало нервическим оживлением каждого мускула; глаза, в которых прежде казался потушенным огонь жизни, теперь блестели лучистым, ярким блеском. Видно было, что чем безжизненнее казался он в обыкновенное время, тем энергичнее был он в эти минуты почти болезненного раздражения.»

Text	Type
<a href="#">Андрей</a>	TAuxDicArticle [наша_первая_грамматика]
<a href="#">Болконский</a>	TAuxDicArticle [наша_первая_грамматика]
<a href="#">Анна</a>	TAuxDicArticle [наша_первая_грамматика]
<a href="#">Павловна</a>	TAuxDicArticle [наша_первая_грамматика]

# Пример неправильной программы

```
S ->Word<h-reg1, fw, wfm=" [А-Я] [а-я, ё] *">;
```

«Князь Андрей, говоря это, был еще менее похож, чем прежде, на того Болконского, который развалившись сидел в креслах Анны Павловны и сквозь зубы, щурясь, говорил французские фразы. Его сухое лицо все дрожало нервическим оживлением каждого мускула; глаза, в которых прежде казался потушенным огонь жизни, теперь блестели лучистым, ярким блеском. Видно было, что чем безжизненнее казался он в обыкновенное время, тем энергичнее был он в эти минуты почти болезненного раздражения.»

Text	Type
<u>князь</u>	TAuxDicArticle [наша_первая_грамматика]
<u>он</u>	TAuxDicArticle [наша_первая_грамматика]
<u>видно</u>	TAuxDicArticle [наша_первая_грамматика]

## Наш проект

Елизавета Петро́вна (18 декабря 1709, Коломенское — 25 декабря 1761, Санкт-Петербург) — российская императрица из династии Романовых с 25 ноября 1741 года, младшая дочь Петра I и Екатерины I, рождённая за два года до их вступления в брак.

