

**Елена Юрьевна Васильева,**  
бакалавр 1-го курса  
Международной школы «Медицина будущего»,  
Института социальных наук  
Первого Московского государственного  
медицинского университета

## **ИЗУЧЕНИЕ НАСЛЕДИЯ ГИППОКРАТА МЕТОДОМ ВЫЧИСЛИТЕЛЬНОЙ СОЦИОЛИНГВИСТИКИ**

В статье представлены результаты исследования истории медицины методом вычислительной социолингвистики. Использован открытый суперкомпьютерный сервис Google Ngram Viewer для работы с электронным корпусом текстов более 30 млн книг на английском языке. Частотный анализ заданных слов позволил обнаружить новые факты в истории медицины и культуры. Самый популярный тип темперамента из наследия Гиппократов — сангвиник. Определены три периода упоминания Гиппократов в англоязычном корпусе оцифрованных книг Google с 1500 года и по наше время.

**Ключевые слова:** Google Ngram Viewer; вычислительная социолингвистика; Гиппократ; сангвиник.

**Elena Vasilyeva,**  
1-year bachelor  
International School «Medicine of the Future»,  
Institute of Social Sciences,  
First Moscow State Medical University

## **STUDY OF HIPPOCRATES HERITAGE BY COMPUTATIONAL SOCIOLOGICAL LINGUISTICS**

The results of research in the fields of History of Medicine and Computational Sociolinguistics are presented in the article. The open supercomputer service Google Ngram Viewer was used to study the digital Corpus of more than 30 million books in English. Frequency analysis of some words allows

discovering the new facts of history and culture. Sanguine is the most popular type of temperament from historical heritage of Hippocrates. Three periods of Hippocrates mention in the English-language Corpus of digitized Google books from 1500 century to present time are defined.

**Keywords:** Google Ngram Viewer; computational sociolinguistics; Hippocrates; sanguine.

Вычислительная лингвистика (математическая обработка языковой информации) существует уже более полувека, ее появление связано с развитием компьютеров и алгоритмов. В настоящее время методы математической обработки электронных корпусов текстов стали называть Text Mining и Text Big Data Analytics. Частотный морфологический анализ (подсчет слов) успешно используется в социологии, например в проектах «Noosphere» [1] или «Третья волна» [2]. Такой вид исследований относят к культуромике или API-социологии, но правильнее относить его к вычислительной социолингвистике [3], так как большой корпус текстов позволяет изучать социум в глобальном масштабе.

В настоящем исследовании использован интеллектуальный суперкомпьютерный открытый сервис Google Ngram Viewer, который создан на основе нескольких технологий вычислительной лингвистики: извлечение данных (Data Mining), распознавание образов (Pattern Recognition), машинное обучение (Machine Learning). Визуализация данных предоставляется по требованию пользователя сервиса в виде гиперактивных графиков (<https://books.google.com/ngrams>).

Сервис Google Ngram Viewer использует корпус оцифрованных книг мирового фонда книгопечатания (более 30 млн книг), начиная с XVI века [4, 5]. Алгоритмы, обрабатывающие по запросу электронный корпус текстов книг, подсчитывают частоту одного заданного слова (процент от всех отдельных слов в текстах за год публикации). Одно слово обозначается как 1-gram.

Проанализированы слова из англоязычного корпуса текстов с 1500 года и по настоящее время. Построены графики подовой динамики частоты слов *sanguine*, *choleric*, *phlegmatic*, *melancholic*, *Hippocrates*, *mortality*. Самый популярный тип темперамента из наследия Гиппократ — сангвиник.

Можно различить три периода в упоминании Гиппократ в корпусе книг, оцифрованных Google: «Рождение» в книгопечатании в середине XVI века; «Золотой век» популярности со второй половины XVII века и до конца XVIII века; «Забвение» начиная с XIX века. «Золотой век» Гиппократ приходится на исторический период борьбы с эпидемиями натуральной оспы. С XV века оспа унесла жизни миллионов европейцев, и только в конце XVIII века появился метод вакцинации (прививка неопасного для человека вируса коровьей оспы), предложенный английским врачом Эдвардом Дженнером (Edward Jenner).

Максимальный пик слова «смертность» (*mortality*) равен максимальному пику слова «Гиппократ» в мировом англоязычном корпусе текстов, но с дистанцией в один век. «Золотой век» Гиппократ XVII века наступает вслед за нарастанием пиков слова «смертность» в XVI веке. Это, вероятно, связано с актуализацией проблемы смертности в данный исторический период и с обращением к учению Гиппократ в условиях борьбы с эпидемиями натуральной оспы — ведущей причины смертности населения Европы в то время. Большой вклад в перевод трудов Гиппократ с древнегреческого языка внес Марк Фабий Кальвусван Равенна (Marcus Fabius Calvusvan Ravenna) в XVI веке. После «Золотого века» Гиппократ прежних пиков слова «смертность» не наблюдается. Таким образом, используя масштабный корпус текстов оцифрованных книг Google, можно проводить социолингвистическое исследование и выявлять новые исторические факты, связанные с восприятием социумом наследия Гиппократ.

### Литература

1. Петров А. М. Доклад «Проект Noosphere. Выявление интересов и поведенческих сценариев пользователей соцмедиа с применением машинного обучения и анализа Больших данных» // VIII Грушинская социол. конф., Секция «Завтрашние технологии социологов». 2018. URL: <https://wciom.ru/index.php?id=1419> (дата обращения: 10.02.2019).

2. Колесниченко О. Ю., Смородин Г. Н., Яковлева Д. А., Мазелис Л. С., Колесниченко Ю. Ю. API-социология: оценка глобальных деструктивных влияний на глокализационную ситуацию // Мониторинг общественного

мнения: Экономические и социальные перемены. 2016. № 6 (136). С. 30–51. DOI: 10.14515/monitoring.2016.6.03.

3. *Nguyen D., Dođruöz A. S., Rosé C. P., de Jong F.* Computational Sociolinguistics: A Survey // *Computational Linguistics*. 2016. Vol. 42 (3). P. 537–593. [https://DOI.org/10.1162/COLI\\_a\\_00258](https://DOI.org/10.1162/COLI_a_00258).

4. Quantitative Analysis of Culture Using Millions Digitized Books / J.-B. Michel, Y. K. Shen, A. A. Presser et al. // *Science*. 2011. Vol. 331 (6014). P. 176–182. DOI: 10.1126/science.1199644.

5. *Эйден Э., Мишель Ж. Б.* Неизведанная территория. Как «Большие данные» помогают раскрывать тайны прошлого и предсказывать будущее нашей культуры / пер. с англ. П. Миронов. (Uncharted: Big Data as a Lens on Human Culture, 2013), М. : АСТ, 2016. 352 с.

УДК 378.147

**Алина Владиславовна Васильева,  
Дарья Евгеньевна Жданова,**  
аспиранты 2-го года обучения  
Уральского гуманитарного института  
Уральского федерального университета

## **ИКТ-ТЕХНОЛОГИИ В ОБУЧЕНИИ ИНОСТРАННЫМ ЯЗЫКАМ**

Психофизиологические особенности цифрового поколения обуславливают поиск новых интерактивных форм обучения ИЯ, включающих в себя современные цифровые ресурсы и увлекательную, наглядную подачу материала. В статье рассмотрены такие современные технологии обучения профессионально ориентированному иностранному языку, как реверсивное обучение и метод проектов, реализация которых невозможна без применения ИКТ.

**Ключевые слова:** ИКТ в обучении, цифровое поколение, метод проектов, смешанное обучение, реверсивное обучение.