

5. Poirot, J. Monte Carlo option pricing for tempered stable (CGMY) processes / J. Poirot, P. Tankov // *Asia-Pacific Financial Market*. – 2006. – Vol. 13, № 4. – P. 327.

6. Кузьмина, А. В. Моделирование дисперсионного гамма-процесса / А. В. Кузьмина // *Вестник национальной академии наук Беларуси. Сер. физико-математических наук*. – 2011. – № 1. – С. 70–74.

УДК [371.016:811.161.'243]:371.31

Линник Л. А., Петросян М. М.

ОБЛАКО СЛОВ КАК МЕТОД КОМПРЕССИИ ИНФОРМАЦИИ НАУЧНОГО ТЕКСТА

Любовь Александровна Линник

старший преподаватель

lyubovlinnik07@gmail.com

Мерри Мгеровна Петросян

старший преподаватель

merryhappy2013@yandex.ru

ФГБОУ ВО «Башкирский государственный медицинский университет»,

Россия, Уфа

TAG CLOUD AS A METHOD OF COMPRESSION OF INFORMATION OF SCIENTIFIC TEXT

Liubov Alexandrovna Linnik

Merri Mgerovna Petrosyan

Bashkir State Medical University, Russia, Ufa

Аннотация. В статье рассматривается один из способов компрессии информации — создание облака слов, который позволяет определить наиболее общие субтемы текста. Даются сервисы, позволяющие составлять облако слов, сам метод компрессии сравнивается с другим методом сжатия научной информации — составлением аннотации. Определяется, что облако, в отличие от аннотации, представляет информацию более обобщенную, в визуальной форме,

аннотация содержит в себе больше фактов, числовых данных, информации о результатах исследования.

***Abstract.** The article discusses one of the methods of information compression - creating a word cloud that allows you to determine the most common subtopics of the text. Services are given that make it possible to compose a word cloud, the compression method itself is compared with another method of compressing scientific information - compiling an annotation. It is determined that the cloud, in contrast to the annotation, represents more generalized information, in visual form, the annotation contains more facts, numerical data, information about the results of the research.*

***Ключевые слова:** облако слов, научный дискурс, компрессия информации, русский язык как иностранный.*

***Keywords:** word cloud, scientific discourse, information compression, Russian as a foreign language.*

21 век, ознаменованный резким скачком объёма ежесекундно поступающей информации, скоростью передачи имеющихся данных и приращением новых, и, конечно, поколением «альфа» среди наших обучающихся, обуславливает, а скорее и вынуждает к нему, поиск новых методов и технологий обучения, трансформации уже традиционных и созданию новых методик. Среди таких методов преподавания русского языка как иностранного можно назвать и программированное обучение, и графические методы (например, создание словообразовательных гнезд на уроках с иностранными обучающимися), и использование формул для организации грамматической информации, и многое другое.

Одним из направлений педагогического и исследовательского поиска стал поиск способа компрессии информации в тексте, любой из методов которой сводится к выделению основной и дополнительной информации, дифференцированию более и менее значимых данных, выведению субтем текста. Среди таких способов компрессии научных текстов можно назвать аннотацию научной статьи, семантическую карту научной статьи и облако слов. Интересно, что аннота-

цию как метаданные любой научной статьи по определённым имеющимся правилам и шаблонам, учитывая институциональность и стандартизованность научного дискурса, а также существующий научный этикет, составляет автор, то есть в ней выделены субтемы текста, наиболее ярко представляющие научную статью с точки зрения автора. Семантическую карту составляет читатель научного текста, выделяя, в свою очередь, важные, с точки зрения актуальности для читателя, субтемы данного текста. Облако слов же составляется программой, искусственным интеллектом, созданным по принципу подражания человеческому разуму, однако при анализе возникает вопрос: чей взгляд на ранжирование информации по принципу значимости в тексте передаёт программа — авторский или читательский. Безусловно, возможно и полное совпадение мнений автора, читателя и искусственного интеллекта, но всё же в большинстве случаев субтемы, выделенные автором, читателем и Облаком, совпадая по многим параметрам, имеют и отличия. В исследовании и обучении представляет интерес как сопоставление авторской, читательской и программной позиций, так и взгляд на каждую из них отдельно. В данной статье мы рассматриваем особенности такого метода компрессии информации, как Облако слов, а также перспективы его использования в учебном процессе.

Облако слов, в другом варианте тегов (tag cloud, word cloud, wordle), представляет собой визуальную модель компрессии информации текста, содержащую лексемы текста, дифференцированные по размеру в зависимости от частотности их употребления. То есть чем чаще слово встречается в тексте, тем больший размер оно принимает.

Рассмотрим несколько способов применения облака слов в исследовательской работе и процессе обучения:

1. Компрессия информации для первичного восприятия текста и его последующего воспроизведения.
2. Анализ текста через определение частотности активных лексем, выделение ключевых слов.
3. Сопоставление текстов посредством облака слов.

4. Сопоставление облака с другими видами репрезентации сжатой информации.
5. Визуальная репрезентация текста как альтернатива диаграмме.
6. Создание инфографики на основе информации, представленной в облаке.
7. Организация и контроль обратной связи (ключевые слова урока, проекта, курса).
8. Мозговой штурм.
9. Моделирование или восстановление содержания текста по облаку слов.
8. Репрезентация отчета и результатов исследования.
9. Определение приоритетных направлений учебных программ или планов.
10. Составление словаря наиболее частотных терминов, функционирующих в научном тексте.
11. Интерполяция текстовых фраз в форме облака в видео.
12. Репрезентация научной информации, лекций, теоретического материала, правил в виде облака (ключевые слова будут выделены и более заметны, чем просто перечень правил).

По сравнению с предшествующими годами популярность «Облака слов» как дидактического инструмента снизилась. Однако использование данного инструмента при анализе научных статей, аннотаций и выделении важных лексем при помощи «Облака» является необходимым, так как оно является наглядным, ясным для большинства и простым в использовании.

Наиболее удобным и интересным сервисом для создания Облака является, на наш взгляд, WordClouds.com, в который можно не только впечатать текст, но и загрузить файлы различных форматов. В WordClouds.com пользователь может создать облака слов любых форм и размеров. Форму слова можно оставить без контура (рисунок 1), а можно добавить контур (рисунок 2).

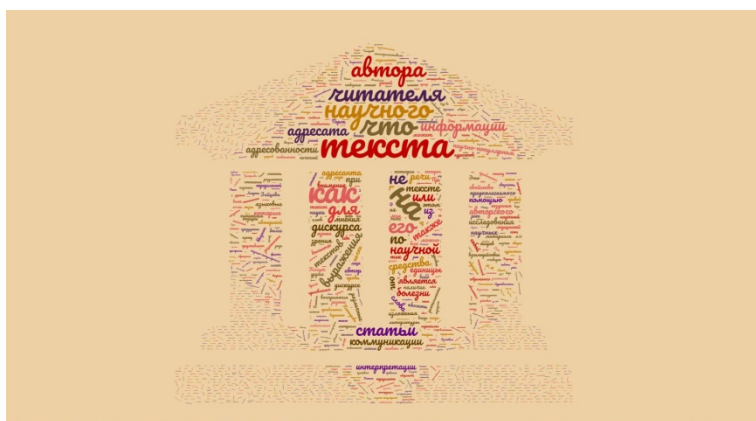


Рисунок 1 — Облако автореферата диссертации канд. филол. Наук Т.М. Михайлюк «Специфика научного текста в аспектах категории адресованности...» [2] без контура



Рисунок 2 — Облако автореферата диссертации канд. филол. Наук Т.М. Михайлюк «Специфика научного текста в аспектах категории адресованности...» [2] с контуром

Кроме того, в данном сервисе доступен широкий спектр цветов, что тоже играет немаловажную роль в наглядности и выживаемости изучаемого материала. Сервис позволяет выбрать расстояние между словами. Также нам доступен выбор размера облака слов. Если какое-то слово не пришлось по душе или наоборот, на ваш взгляд, здесь какого-то слова не достаёт, то данный нюанс тоже легко исправить. Такие возможности корректируют вид облака слов в сторону взгляда на данный текст более читателя, нежели автора. Данный сервис позволяет сохранить файл в форматах JPG, PDF, PNG. Кроме того, файлом можно сразу же поделиться в социальных сетях.

Следующий сервис, который мы рассмотрим, — это WordWanderer. Несмотря на оригинальность данной платформы, которая заключается в том, что пользователь сам может менять слова местами, делая ядро таким, каким ему оно

видится, имеется ряд существенных недостатков. Во-первых, потому что данный сервис не даёт пользователю возможность скачать получившееся облако слов. Выйти из данной ситуации можно, сделав скриншот, однако он, видимо, рассчитан лишь на одномоментное онлайн использование. Кроме того, у данной платформы очень бедный интерфейс: нельзя выбрать или изменить цвета (черно-белый), нет возможности выбрать размер картинки, размер шрифта, форму облака слов. Загрузить текст для создания интеллектуальной карты можно только посредством копирования, загрузить файл тоже не представляется возможным. Убрать те или иные части речи из облака слов также не получится, так как оно, по-видимому, не рассчитано на русский язык. Несмотря на существование такой функции, применить её так и не получилось (рисунок 3).

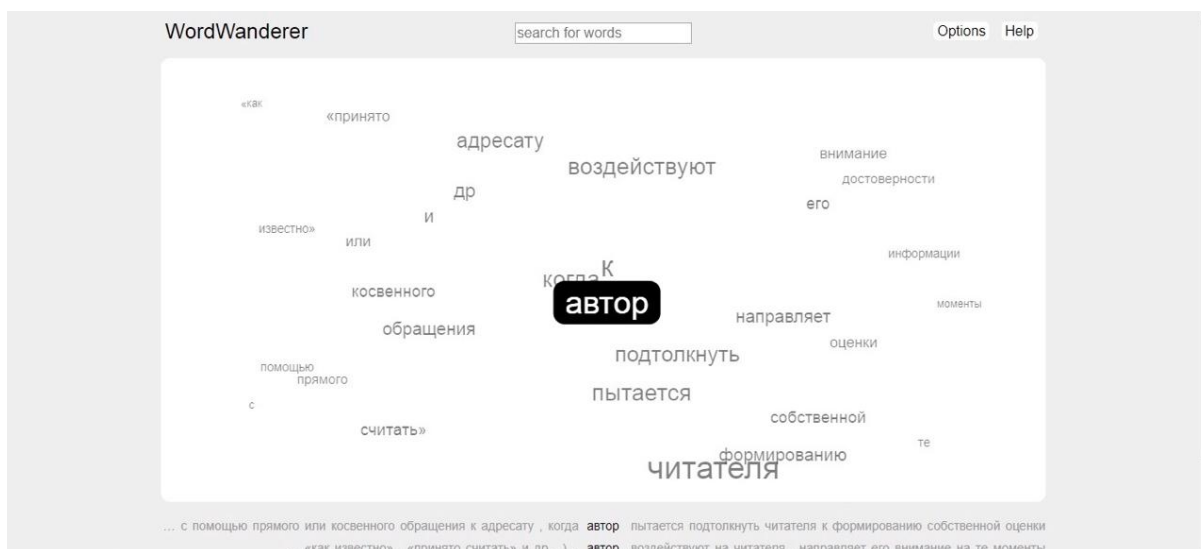


Рисунок 3 — Облако автореферата диссертации канд. филол. наук Т. М. Михайлюк «Специфика научного текста в аспектах категории адресованности...» [2] **WordWanderer**

Ещё один сервис, позволяющий создать облако слов какого-либо текста, Облакослов.рф. Этот онлайн-генератор позволяет вводить весь текст методом копирования, выбирать палитру, регулировать количество слов для отображения в облаке, фон, шрифт, размер изображения, минимальную частоту слов для отображения в облаке, ссылку на источник текста, обозначить стоп-слова, то есть слова, которые не будут отображаться в тексте (например, несамостоятельные части речи), а также дает возможность вместо текста ввести ссылку на RSS-feed веб-сайта вместо текста, что снимает ограничения по объёму и размеру текста,

по которому создаётся облако. Язык текста может быть любым. Этот сервис представляется нам более нейтральным по отношению к позициям автора и читателя, поскольку создатель облака может вносить изменения больше в его форму, нежели в содержания. Предоставляемая сервисом возможность введения стоп-слов существенно на выделение субтем не влияет, так как стоп-словами, как правило, являются лексемы, не содержащие значимую информацию, а использующиеся для связи слов и предложений.

Облакослов.рф предоставляет образцы наиболее популярных облаков слов, какими на момент написания статьи, например, являлись Послание Путина Федеральному Собранию 2020 (главная субтема — «Мы»), новогоднее обращение Путина (главная субтема также — «Мы»), Мюнхенская речь Путина (главная субтема — «Безопасность»), отзывы на iPhone 11 (главная субтема — «Репутация») и др.

Применение технологии составления облака слов позволяет представить сжатую информацию научной статьи в наглядной форме так, что наиболее значимые субтемы сразу определяются читателем. Так, например, статья «Артериальная гипертензия: распространенность, осведомленность, прием антигипертензивных препаратов и эффективность лечения среди населения российской федерации» Шальной С. А., Балановой Ю. А., Константинова В. В., Тимофеевой Т. Н., Иванова В. М., Капустиной А. В., Деева А. Д. [3] в облаке выглядит следующим образом:

наличии АГ — 77,9 % (выше среди **женщин** — 80,3 %, против 75 % у **мужчин**). Лечатся 59,4 % **населения** (лучше — **женщины** — 63,1 % против 53,1 % среди **мужчин**), причем эффективно 21,5 % (22,5 % **женщин** и 20,5 % **мужчин**). На первое место по числу назначений вышли ингибиторы АПФ (ИАПФ) — 70,7 %.

Ключевые слова: дескриптивная эпидемиология, **артериальная гипертония, распространенность, осведомленность, эффективность терапии, ингибиторы АПФ, федеральные округа.**

Как мы видим, темы, выделенные искусственным интеллектом, точно совпадают с темами, определенными авторами статьи, различие двух методов компрессии информации заключается в большей приверженности научному дискурсу авторами, выражающейся в реализации стратегий информирования и обобщения.

Таким образом, составление облака слов как способ компрессии информации и метод обучения позволяет получить основную наиболее значимую информацию о содержании научного текста, исключая при этом необходимую для эффективности научного дискурса числовую и фактическую информацию, а также когнитивные авторские стратегии. Использовать облако можно для создания общего понимания тематики текста и определения, интересен ли он для дальнейшего изучения.

Список литературы

1. Линник, Л. А. Количественные методы в преподавании русского языка как иностранного: дифференциация и эффективность / Л. А. Линник, М. М. Петросян // Довузовский этап обучения в России и мире: язык, адаптация, социум, специальность / ответственные редакторы: М. Н. Русецкая, Е. В. Колтакова. Москва, 2017. – С. 357–361.

2. Михайлюк, Т. М. Специфика научного текста в аспектах категории адресованности: (Сопоставительный анализ французских и русских научных текстов): автореф. дис... канд. филол. наук: 10.02.01 / Татьяна Михайлова Михайлюк. – Барнаул, 1996.

3. Шальнова, С. А. Артериальная гипертония: распространенность, осведомленность, прием антигипертензивных препаратов и эффективность лечения среди населения российской федерации / С. А. Шальнова, Ю. А. Баланова, В. В. Константинов, Т. Н. Тимофеева, В. М. Иванов, А. В. Капустина, А. Д. Деев // Российский кардиологический журнал. – 2006. – № (4) – С. 45–50. – URL: <https://doi.org/10.15829/1560-4071-2006-4-45-50>.

УДК 61:001.891.5:004

Мещанинов В. Н., Гаврилов И. В.

**АНАЛИЗ ЦИФРОВЫХ БАЗ ДАННЫХ БИОЛОГИЧЕСКОГО ВОЗРАСТА
В МЕДИЦИНСКОМ ОБРАЗОВАНИИ**

Виктор Николаевич Мещанинов

Доктор медицинских наук профессор заведующий кафедрой биохимии

mv-02@yandex.ru

Уральский государственный медицинский университет, Россия, Екатеринбург

Институт медицинских клеточных технологий, Россия, Екатеринбург,

Илья Валерьевич Гаврилов

Кандидат биологических наук доцент кафедры биохимии

iliagavrilov18@yandex.ru

Уральский государственный медицинский университет, Россия, Екатеринбург

Институт медицинских клеточных технологий, Россия, Екатеринбург

**ANALYSIS OF DIGITAL DATABASES OF BIOLOGICAL AGE IN
MEDICAL EDUCATION**

Viktor Nikolaevich Meshchaninov

Ural state medical University, Russia, Yekaterinburg

Institute of Medical Cellular Technologies, Russia, Yekaterinburg

Iliia Valerevtsch Gavrillov

Ural state medical University, Russia, Yekaterinburg

Institute of Medical Cellular Technologies, Russia, Yekaterinburg