

---

---

## ОПЫТ ПРАКТИЧЕСКОГО ПРИМЕНЕНИЯ СЕТЕВОГО АНАЛИЗА



DOI: 10.19181/4m.2023.32.1.4

EDN: FDYTSV

С. Ткач, П.Д. Воробьева, М.М. Русакова  
(Санкт-Петербург)

### ОПЫТ РЕАЛИЗАЦИИ ДИСКУРС-АНАЛИЗА И КОНЦЕПТУАЛЬНОГО КАРТИРОВАНИЯ СООБЩЕСТВ ЗДОРОВОГО ПИТАНИЯ<sup>1</sup>

В статье приводится опыт имплементации методов дискурс-анализа в трактовке Э. Лакло и Ш. Муфф и метода концептуального картирования в интерпретации У. Троича посредством методики сетевого анализа на примере тематики здорового питания. Результатом анализа выступают графы, которые позволяют выделить борьбу агентов дискурса за значение ключевых дискурсивных знаков, а также рассмотренная в статье в качестве примера концептуальная карта участников онлайн-дискуссии по спорным в рамках тематики вопросам. В качестве эмпирической базы дискурс-анализа выступили 3 тыс. собранных комментариев в четырех сообществах о здоровом питании социальных сетей «ВКонтакте» и «Одноклассники». Разработанная реализация методов концептуального картирования и дискурс-анализа адаптирована для анализа онлайн-дискуссий. Последующая валидация разработанных методов видится

---

**Сергей Ткач** – социолог, Центр прикладной социологии, Санкт-Петербургский государственный университет, Санкт-Петербург, Россия. Email: s.tkach@spbu.ru.

**Полина Дмитриевна Воробьева** – социолог, Центр прикладной социологии, Санкт-Петербургский государственный университет, Санкт-Петербург, Россия. Email: st098355@student.spbu.ru.

**Майя Михайловна Русакова** – кандидат социологических наук, директор, Центр прикладной социологии, Санкт-Петербургский государственный университет, Санкт-Петербург, Россия. Email: m.rusakova@spbu.ru.

направлением дальнейших исследований. У предлагаемых дизайнов есть ряд ограничений, которые обсуждаются в статье.

*Ключевые слова:* дискурс-анализ, концептуальное картирование, анализ графов, сентимент-анализ, анализ естественного языка

## *Введение*

Социальные сети обладают рядом особенностей, которые затрудняют их исследование. Структура интерфейса социальных сетей может значительно влиять на то, как происходит дискуссия [1; 2], и делать ее отличной от офлайн-дискуссий или дискуссий в печатных и телевизионных СМИ. Это требует от исследователей разработки методов, более подходящих под формат онлайн-дискуссии. Возможное решение видится в совершенствовании существующих исследовательских методов посредством имплементации цифровых элементов анализа. Методы дискурс-анализа и концептуального картирования имеют длинную традицию эффективного применения учеными для решения исследовательских задач анализа текста. Однако в своем изначальном дизайне они не ориентированы на анализ дискуссий в социальных сетях. Имплементация цифровых элементов анализа позволит ухватить данную специфику. Также комбинация методов позволит получить более детальное представление об исследуемом объекте. Реализация этих методов будет продемонстрирована на примере дискуссий в онлайн-сообществах здорового питания.

## *Концептуальное картирование*

Концептуальное картирование было разработано как самостоятельный метод У. Трочимом в начале 1980-х гг. Метод совмещал в себе элементы качественного анализа (мозговой штурм, неструктурированная сортировка, интерпретация) и количественного

анализа (многомерное шкалирование, иерархический кластерный анализ). Метод позволяет группе людей изобразить наглядную последовательную концептуальную схему интересующего их вопроса [3]. Сам Трочим определяет концептуальное картирование как структурированный процесс, сосредоточенный на интересующей теме, включающий вклад нескольких участников, итогом которого становится интерпретируемое графическое представление идей и концепций участников построения, а также то, как эти идеи взаимосвязаны. Под концептуальной картой Трочим понимает структурную концептуализацию или многомерную графическую репрезентацию набора идей, сгенерированных группой [4]. Изначально метод включал в себя шесть этапов: подготовка к проведению исследования (включая отбор участников дискуссии); производство высказываний участниками дискуссии; структурирование полученных высказываний; представление высказываний в виде концептуальной карты (с использованием многомерного шкалирования и кластерного анализа – сам Трочим рекомендовал для удобства ограничиться двумерным представлением); интерпретация полученных карт; практическое использование карт в решении прикладных задач [5]. Шестой этап является прикладным, поэтому исследователи, как правило, выделяют пять аналитических этапов. За прошедшие десятилетия метод приобрел множество вариаций и интерпретаций в рамках исследований в общественных науках. В частности, можно упомянуть работу Н. Абрамовой и Ю. Николаевой об использовании концептуального картирования как метода повышения валидности результатов оценочного исследования [6], а также шестикомпонентную модель реализации метода, разработанную Трочимом совместно с М. Кейном [7]. Метод концептуального картирования предполагает работу с офлайн-дискуссиями специально отобранных экспертов. Однако он показал свою эффективность и для анализа групповых дискуссий экспертов опыта [8] (информантов, обладающих экспертностью, приобретенной посредством проживания уникального опыта,

а не наличием квалификации [9]), что делает перспективным использование этого метода для анализа онлайн-дискуссий. Однако сама специфика онлайн-дискуссий требует доработки метода посредством имплементации элементов анализа цифровых следов. Авторы в статье опираются на модель Трочима «концептуальной карты» и общую логику его подхода, внося изменения в метод концептуального картирования таким образом, чтобы он стал более подходящим для анализа онлайн-дискуссий. Мы понимаем под концептуальной картой интерпретируемое графическое представление сообщений участников онлайн-дискуссии, а также то, как идеи, изложенные в сообщениях, взаимосвязаны между собой.

### *Дискурс-анализ: концептуализация*

В академической литературе не существует единого мнения о сущности понятия «дискурс» и, как следствие, о понятии «дискурс-анализ». В понимании Э. Лакло и Ш. Муфф дискурсом можно назвать упорядоченную тотальность знаков. Лакло и Муфф полагают, что некоторые знаки более ценны, чем другие. К таким знакам обращено больше внимания, их чаще изображают, о них чаще говорят, от их значения зависят значения других знаков, менее ценных. Их можно назвать *узловыми точками* – они сплетают воедино дискурс, фиксируют его. Содержание узловых точек пусто и изменчиво: разные дискурсы могут наделять его значением, тем самым обозначив свое превосходство над другими дискурсами. Здесь появляется другое ключевое понятие теории Лакло и Муфф – *борьба*. Различные дискурсы постоянно *ведут борьбу*, стремясь зафиксировать свое значение в языке. Отсюда становится понятна одна из основных целей дискурс-анализа в предложенной трактовке: проследить процессы борьбы за определенный способ определения значений в языке [10].

Новым методологическим вызовом для дискурс-анализа становится развитие цифровых технологий. Несмотря на то, что существу-

ющее в теории Лакло и Муфф всеобъемлющее постмодернистское понятие дискурса позволяет обращаться к различным формам данных в качестве объектов исследования, аналитическое переложение теоретических концепций исходной теории к интернет-дискуссиям – достаточно нетривиальная задача [11]. Дискурс-анализ в осмыслении Лакло и Муф – это метод, используемый для анализа политического дискурса, и возможности применения аналитического аппарата этого метода для исследования других сфер общественной жизни неочевидны. Так, потенциал использования метода дискурс-анализа для исследований медиа показали Н. Карпентер и Б. Де Клине [12]. В их совместной статье в том числе рассматривается возможность изучения медиа в качестве поля дискурсивной борьбы. О. Игнатевой отмечается, что сама специфика интернет-дискуссий, такая как особый интернет-язык, для которого характерны сокращения или киберорфография, усложняет процесс аналитической обработки дискурса. В дополнение интернет-дискуссии могут быть рассмотрены в качестве устной речи, что требует особых методов анализа, например – конверсационного анализа [13]. Большое число исследований, направленных на изучение большого количества текстовых данных и предполагающих цифровые этапы обработки дискурса, как отмечает Дж. Хадитаги и другие ученые, имеют существенный недостаток в виде назначения узловых точек «внешней силой», что может не соответствовать эмпирическому объекту изучения [14]. В статье П. Бакумова приводятся примеры нескольких исследований, в рамках которых в качестве инструментов анализа использовалась цифровая имплементация метода дискурс-анализа в трактовке Лакло и Муфф. Бакумов отмечает, что большинство попыток использования выбранного метода дискурс-анализа начинаются с переопределения теоретических концепций исходного метода, а также последовательности шагов анализа [15]. Однако не сложно заметить, что каждая имплементация этого метода крайне специфична и напрямую зависит от предмета и целей исследования. Это к приводит к выводу, что проведение дальнейших

исследований в сфере использования цифровых методов анализа дискурса необходимо.

Авторы обращаются к традиции критического дискурс-анализа в интерпретации Лакло и Муфф. Ключевым изменением в сравнении с традиционным дискурс-анализом является использование методов визуализации, посредством которых наглядно изображается встречаемость тех или иных знаков дискурса. В понимании Лакло и Муфф поле дискурса – это поле борьбы. Значение, скрывающееся за знаками, не постоянно – оно устанавливается различными агентами, преследующими свои интересы посредством высказываний – дискурсивных практик. Интересанты, чью позицию выражают агенты дискурса, в большинстве случаев не могут быть напрямую соотнесены с конкретными людьми. В повседневной речи люди упоминают слова с разным контекстуальным значением, транслируя тем самым разные интересы, скрытые за этим. Агент (действующее лицо или группа) производит высказывание, в котором упоминает какую-то *узловую точку*. Делает он это в определенном контексте: в высказывании упоминается не только сама узловая точка, но и другие знаки. Этот контекст (множество знаков, упоминаемое в высказывании) способен, согласно Лакло и Муфф, переопределить значение *узловой точки*. Так как этот контекст разный у разных агентов, значение *узловой точки* постоянно переопределяется. Те агенты, кто делают это успешнее, побеждают на поле боя дискурса. Победа может быть с некоторыми допущениями названа конечной, когда определенное контекстуальное толкование слова полностью исчезает из речи. Такую окончательную победу, как и промежуточные, можно зафиксировать в рамках дискурс-анализа.

## *Дискурс-анализ: методология*

### *Методология отбора данных*

Исследовательской группой были отобраны четыре сообщества социальных сетей «ВКонтакте» и «Одноклассники», посвя-

щенные здоровому питанию<sup>1</sup>. Отбор комментариев проходил с 16 апреля по 6 мая 2023 года.

Сообщества были отобраны по следующим критериям.

1. Количество подписчиков: отбирались наиболее популярные сообщества. Ранжирование происходило при помощи поиска в сетях «ВКонтакте» и «Одноклассники».

2. Активность сообщества. Если сообщество было неактивным, то есть в нем не было ни одного поста за последнюю неделю, оно удалялось. По итогам выборки ни одно сообщество не было удалено.

Из отобранных сообществ были агрегированы все комментарии, когда-либо оставленные пользователями, – 15 618 комментариев. К комментариям сохранялись также метаданные: пост, под которым они были размещены, пользователи, оставившие их, время комментария, количество отметок «нравится». Критерии для отбора 3000 комментариев для анализа указаны ниже.

### Методология анализа

Главной метафорой, вокруг которой выстраивалась разработка метода – поле дискурса, представляющее собой систему упорядоченных знаков посредством связей встречаемости. Связь и борьба,

---

<sup>1</sup> Работа выполнена при поддержке СПбГУ, шифр проекта 121062300141-5. Исследование, в котором были реализованы концептуальное картирование и дискурс-анализ, было посвящено различным потребительским практикам россиян в зависимости от их уровня финансового благополучия.

Анализируемые сообщества:

- 1) «КЕТО диета», ссылка на сообщество «ВКонтакте»: <https://vk.com/ketodieta>;
- 2) «Диета Углеводов.нет», ссылка на сообщество «ВКонтакте»: <https://vk.com/uglevodovnet>;
- 3) Greenway Global, ссылка на сообщество «ВКонтакте»: <https://vk.com/greenwayglobalofficial>;
- 4) «Павел Корпачев – здоровье, фитнес, тренировки», ссылка на сообщество в «Одноклассниках»: <https://ok.ru/zenofit>.

которые подразумевались Э. Лакло и Ш. Муфф иносказательно, могут быть изображены наглядно. Встречаемость (контекстность) одного знака по отношению к другому может быть представлена в виде ребра графа. В исследовании под знаком операционально понимается слово. Тогда знаки с наибольшим числом ребер становятся узловыми точками дискурса. Под агентами понимаются пользователи сообщества. То, какие агенты используют знаки, можно выделить цветом. Так, словосочетания, которые используются агентами с одной позицией, будут окрашены в один цвет, а агентами с другой позицией – в другой. Под позицией понимается ценностная установка, которая кажется агентам правильной и которую они стараются донести в своих высказываниях. Борьба агентов сводится к образованию как можно большего числа ребер с узловой точкой, другими словами – к расширению пространства контекста точки дискурсивной борьбы. Теперь перейдем непосредственно к шагам реализации такой модели.

Первым шагом была кодировка комментариев. Исследовательская группа знакомилась с подвыборкой из 1000 комментариев. Данная подвыборка состояла из комментариев, которые получили наибольшее количество отметок «нравится». Наибольшее число отметок «нравится» составило 180, а наименьшее – 4. Среди всех комментариев с 4 отметками «нравится» были отобраны самые недавние – как представляющие наиболее актуальный дискурс в сообществе. Размер выборочной совокупности ограничивался трудовыми ресурсами исследовательской группы.

**Процедура кодировки.** Прочитав комментарии, группа выделяла позиции агентов дискурса (классы) (например, «худеющие» – те, кто придерживается позиции, что здоровое питание – это инструмент в процессе похудения и т.д.) и присваивала каждому комментарию соответствующую метку – от 1 до n (при n выделенных позиций). Описание позиций агентов записывалось. Ручная кодировка комментариев, получивших наибольшее количество оценок «нравится», обусловлена тем, что позволила выделить те группы

агентов, которые наиболее влиятельны в исходном дискурсе. На первом этапе были сформированы и уточнены инструкции кодирования, на втором этапе эти инструкции использовались дополнительными кодировщиками. Исследовательской группой были вручную классифицированы в соответствии с выделенными при кодировке позициями другие случайно отобранные 2000 комментариев. Случайный отбор осуществлен функцией `sample` библиотеки `Pandas`.

Следующим пунктом была подготовка текста комментариев для дальнейшего анализа. В предобработке использовались библиотеки языка `Python`: `rumorphy2` и `NLTK`:

- 1) была удалена вся пунктуация в тексте;
- 2) слова были приведены к нормальной форме и нижнему регистру;
- 3) среди слов были отобраны только прилагательные и существительные;
- 4) все слова в тексте были разбиты на словосочетания (биграмы методом `nlk.bigrams`) с указанием частоты их встречаемости;
- 5) из полученных словосочетаний были дополнительно удалены те, которые содержали наиболее общеупотребимые слова – слова бытовой повседневной речи (при помощи словаря библиотеки `NLTK` языка `Python`);
- 6) из словосочетаний был построен граф при помощи программы `Gephi` – ребрами графа выступала совстречаемость слов на расстоянии 1 в тексте.

## *Результаты*

Граф поля дискурса строился из словосочетаний, приведем наиболее популярные из них (рис. 1).

Основная часть посвящена ведению диет: день кето, неделя кето, неделя два, вернуться кето – эти словосочетания посвящены кетодиете. Другие популярные словосочетания посвящены прак-

тикам приема пищи, и в особенности их частоте и продолжительности: прием пищи, каждый день, раз [в] день, первый день, течение день, пара день [дней], каждый прием, прием неделя. Третья группа наиболее популярных словосочетаний включает в себя пищевые продукты и соединения: сливочный масло, яблочный уксус, фульвовая кислота, цитрат магний. В последнюю группу

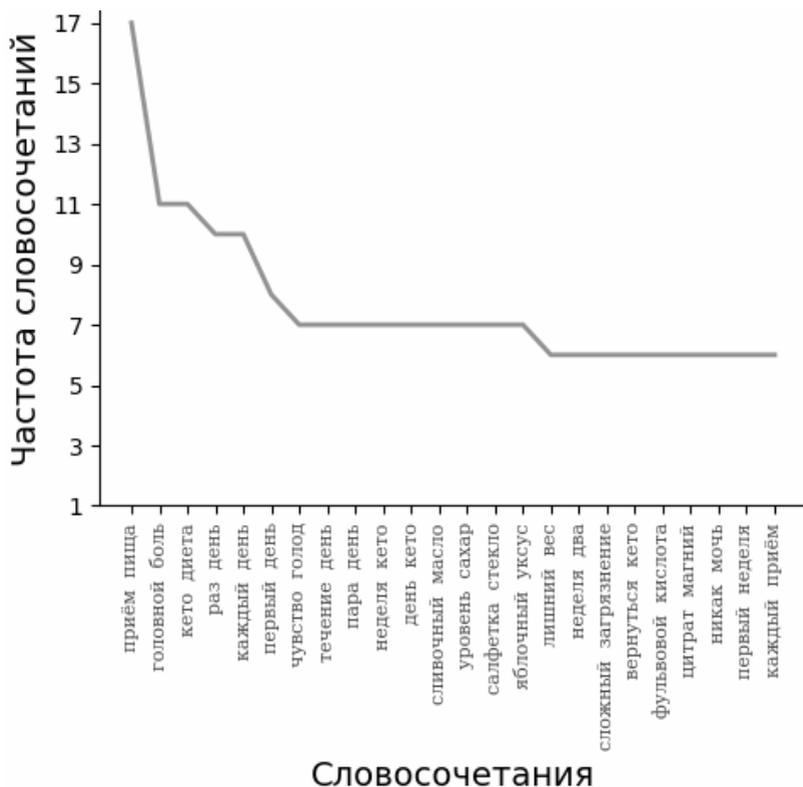
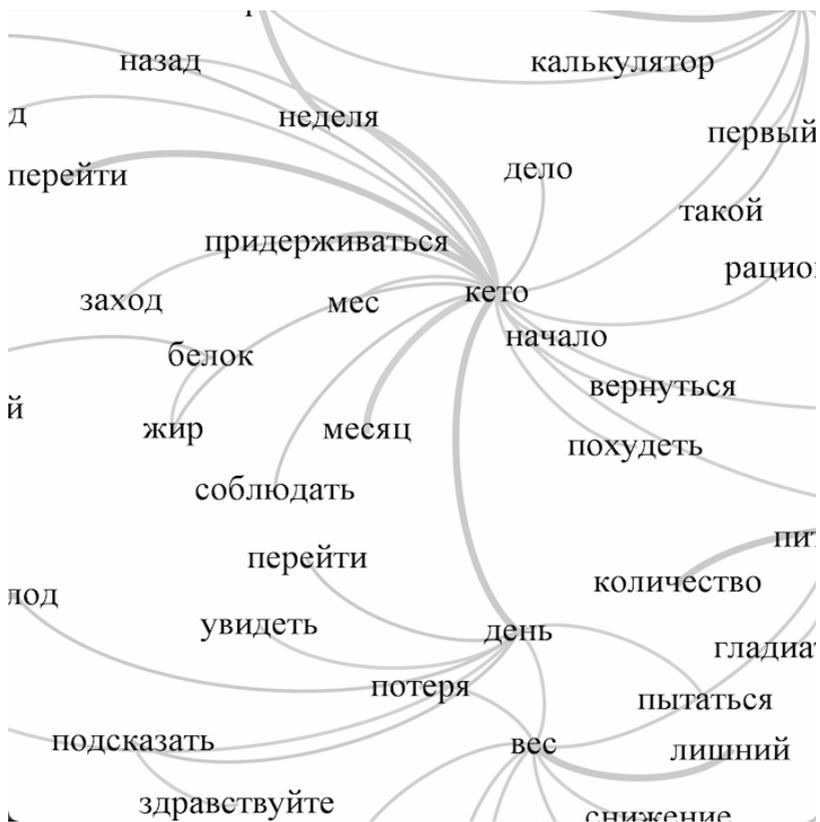


Рис. 1. Частотное распределение наиболее популярных словосочетаний сообществ здорового питания

можно отнести ментальные и соматические состояния: головной боль, чувство голод, уровень сахар, лишний вес, никак [не] мочь.

Число узлов составило 648, а связей – 440 (рис. 2). Плотность полученного графа составила 0,002, средняя степень центральности – 1,358, средняя длина пути в графе – 3,916, средняя степень близости узла – 0,76636, модулярность – 0,94, число обнаруженных компонент методом Гирвана-Ньюмана составило 222.



*Рис. 2. Фрагмент поля дискурса здорового питания*

## *Дискуссия*

В рамках описанной в методологии кодировки были выделены шесть групп агентов дискурса.

1. Худеющие – агенты, которые делятся результатами своего похудения. Они воспринимают здоровое питание как способ эффективно похудеть, привести себя в форму (10,62% комментариев, здесь и далее – от общего объема выборочной совокупности в 3000 комментариев).

2. Кулинары – агенты, которые с интересом делятся рецептами и рационами. Для них здоровое питание – это возможность вкусно приготовить и поесть здоровую пищу (19,01% комментариев).

3. Комментаторы – агенты, которые реагируют на публикации сообщества, критикуют, восхищаются и др. В их представлении здоровое питание – это бренд, организация, собеседник (19,01% комментариев).

4. Лайфстайл – агенты, обсуждающие обычные повседневные привычки и радости. Практика здорового питания не имеет ключевого значения в их речи (17,53% комментариев).

5. Опытные – агенты, которым важно поделиться опытом в освоении здорового питания, возникшими ранее проблемами и путями их решения. Для них здоровое питание уже было важной частью жизни (15,8% комментариев).

6. Члены комьюнити – агенты, активно участвующие в общении с другими членами сообщества, дающие советы, задающие вопросы, поддерживающие других. В их представлении здоровое питание – это общее дело, возможность найти единомышленников или обогатиться новыми знаниями (18,03% комментариев).

Агенты образуют поле данного дискурса (рис. 2). В качестве точек дискурсивной борьбы выступают знаки «кетто», «день», «питание», «вес». Эти знаки содержат наибольшее число связей с другими – подсчет производился членами исследовательской группы. Результаты анализа приводятся в табл. – для компактно-

*Таблица*

ПРОАНАЛИЗИРОВАННЫЕ ЗНАКИ НА ПОЛЕ ДИСКУРСА

Знак	Встречаемые совместно знаки	Агенты, использующие встречаемые совместно знаки
Кето	Неделя, месяц, год, вернуться, заход, придерживаться, похуделый, похудеть	Опытные
	Пропорция, жир, белок, питание	Члены комьюнити
День	Первый, третий, пара, следующий, раз, вес	Опытные
	Добрый, подсказать	Члены комьюнити
	Пытаться, нагрузка	Лайфтсайл
Питание	Скидка, спортивный	Лайфстайл
	Начинать, низкоуглеводный, кето	Члены комьюнити
Вес	Потеря, держать, терять	Худеющие
	Встать, потеря	Опытные

сти приведены только наиболее многочисленно представленные позиции агентов.

Точкой наиболее ожесточенной дискурсивной борьбы оказывается знак «кето» – его используют 5 классов агентов из 6. Рассмотрим только наиболее активные группы агентов в этой борьбе. «Опытные» используют знак «кето» совместно со знаками временных рамок (неделя; месяц; год), а также в контексте возвращения и повторяющейся практики (вернуться; заход; придерживаться и др.). Они также указывают на эффективность диеты в похудении (похуделый; похудеть). «Члены комьюнити» используют знаки для уточнения деталей и нюансов практики «кето» (пропорция; белок; жир; питание), пытаются точно сформулировать стратегию здорового питания. Итогом борьбы за знак «кето» становится выигрыш «опытных» агентов – знак «кето» обладает наибольшим числом контекстуальных связей в речи «опытных» агентов.

Следующей точкой дискурсивной борьбы является «день». Если агенты класса «опытные» используют данное слово в контексте своего прошлого опыта, а также для обозначения распорядка практики здорового питания (первый; третий; пара; следующий; раз; вес), то «члены комьюнити» больше фокусируются на доброжелательном тоне общения и взаимной поддержке (добрый [день]; подсказать и др.). Для представителей класса «лайфстайл» соблюдение практики здорового питания является тяжелой задачей (пытаться; нагрузка и др.). Победу в данной точке дискурсивной борьбы вновь одерживают «опытные» агенты.

Следующей точкой дискурсивной борьбы является знак «питание». Активнее всего борьбу за данный знак ведут классы агентов «лайфстайл» и «члены комьюнити». Для представителей класса «лайфстайл», важен разумный экономический подход (используя знак «скидка»), они также пытаются связать с практикой здорового питания другие сферы своего интереса (спортивный). Говоря о питании, «члены комьюнити» ищут единомышленников на начальных этапах практики здорового питания (начинать; низкоуглеводный; кето). Именно этот класс побеждает в борьбе за знак «питание».

Последней точкой дискурсивной борьбы является знак «вес». За него борются классы «худеющих» и «опытных». Другие классы участвуют в борьбе несущественно. Для класса «худеющих» «вес» фигурирует в контексте достижения основной цели – похудения (потеря; держать; терять). Агенты из класса «опытных» используют знаки «стоять», «встать» наряду со словом «снижение», что говорит о достижении цели и желании сохранить результат. В результате борьбы данный знак остается за классом «худеющих» агентов.

## *Концептуальное картирование*

### *Методология отбора данных*

Для построения концептуальной карты был отобран пост в сообществе о здоровом питании, под которым была дискуссия, содержащая наибольшее число комментариев для данного сообщества, – 19. Данная дискуссия была выбрана в качестве примера для реализации концептуального картирования. На ее основе создавался свой ориентированный граф – концептуальная карта. Данные, с помощью которых была получена концептуальная карта, содержали следующие характеристики: текст комментария, автор комментария, время его публикации и количество выставленных сообщению лайков.

### *Методология анализа*

Процесс построения концептуальной карты можно разделить на несколько этапов. Первым этапом построения является первичная обработка сообщений – естественного языка, – направленная на упрощение текста для последующей машинной обработки. Также анализировалась тональность высказывания. Текст был подвергнут минимальной первичной обработке – приведен к нижнему регистру, очищен от эмодзи и специальных символов. После описанной предобработки из полученного массива данных удалялись пустые комментарии.

Вторым этапом реализации является определение семантической близости между высказываниями участников дискуссии. Определение семантической близости текстов – широко распространенная задача в области обработки естественного языка [16] и включает в себе две подзадачи: векторное представление слов и непосредственно определение семантической близости. В результате векторного представления слов анализируемый текст трансформируется в вектор. Это позволяет применять к анализу

текстов методы машинного обучения. Векторное представление слов осуществляется с помощью библиотеки `sentence_transformers`. Выбранный из библиотеки метод `SentenceTransformer` реализуется на базе предобученной мультиязычной нейронной сети `Sentence-BERT (SBERT)`, позволяющей корректно сравнивать русскоязычные комментарии [17]. После векторной трансформации текстов комментариев необходимо определить семантическую близость каждой пары комментариев. Для этого создается матрица, заполненная нулями, размерности  $N \times N$ , где  $N$  – количество комментариев в дискуссии. К каждой паре комментариев применяется метод `cosine_similarity` из библиотеки `sklearn.metrics.pairwise` для расчета значений косинусного расстояния. Семантическая близость между комментариями, таким образом, равна косинусному расстоянию между двумя векторами, образованными при помощи векторного представления этих комментариев нейронной сетью `Sentence-BERT`.

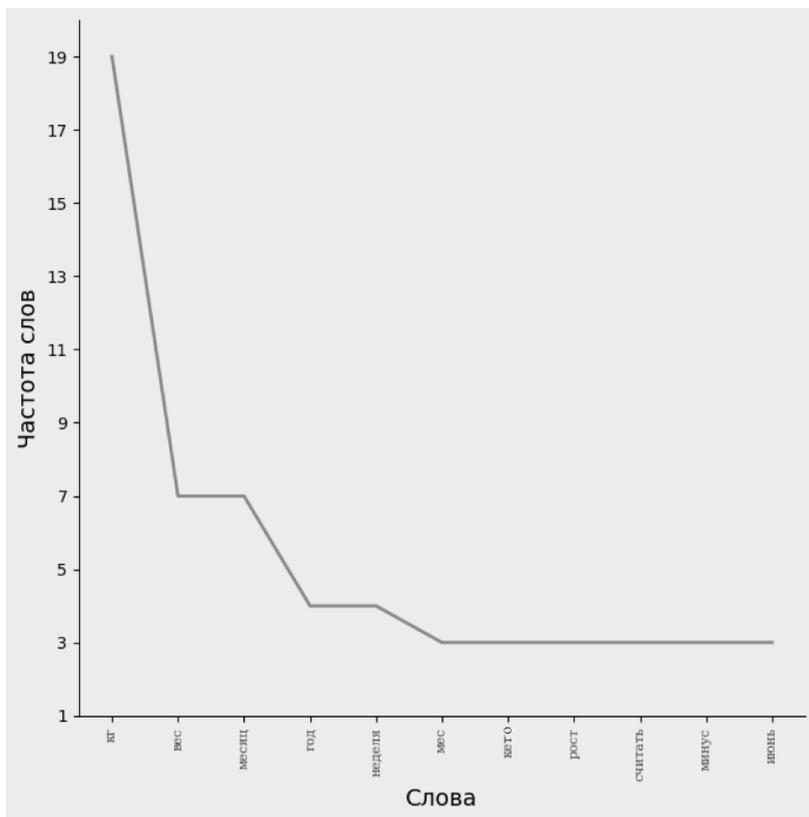
На третьем этапе определяется эмоциональная окраска комментариев. Данная информация была необходима для определения логической связи между сообщениями дискуссии. Если негативно окрашенный комментарий был семантически близок к комментарию с положительной тональностью, предполагалось, что логическая связь между комментариями строго дизъюнктивна. Нейтральная связь предполагала только наличие причинной связи и т.д. Эмоциональная окраска комментариев дискуссии определялась с помощью библиотеки языка Python `dostoevsky`. Ее особенностью является обучение модели на корпусе русскоязычных текстов [18]. Модель определяет вероятность принадлежности каждого комментария к одному из пяти классов тональности: `neutral` (или нейтральный тон), `negative` (негативный), `positive` (положительный), а также `speech` (текст является элементом разговорного языка и не содержит ярко выраженной эмоциональной окраски), `skip` (данный класс присваивается тексту, если модель не может определить тональность). В рамках построения концеп-

туальных карт с помощью модели каждому комментарию присваивался наиболее вероятный класс из перечисленных выше. Итогом данного этапа является перечень меток эмоциональной окраски сообщений, информация о которых добавляется в изначальный массив комментариев.

Наконец, четвертый этап: на основе исходного массива комментариев, с добавленной информацией об их тональности, а также сконструированной на предыдущем этапе матрицы косинусных расстояний, создается ориентированный граф. В силу ограниченности трудовых ресурсов экспертной группы проекта значение косинусного расстояния, при котором в итоговый ориентированный граф добавляются связи между вершинами, подбирался вручную таким образом, чтобы граф имел не более 20 ребер между вершинами-сообщениями – общего числа ребер графа. При увеличении числа ребер количество анализируемых цепочек комментариев многократно росло, отчего было принято решение ограничиться наиболее связанными комментариями – ядром дискуссии. Такой граф представляет собой концептуальную карту дискуссии, в которой отражен ход проанализированной дискуссии. Его визуальное представление строится посредством программы Gephi. Вершинами графа выступали комментарии дискуссии, а ребра между ними обозначались, если семантически комментарии были достаточно близки друг к другу. Ребра были направлены от более ранних комментариев к более новым, что позволяло сохранить причинно-следственные связи между сообщениями.

## *Результаты*

В дискуссии явно выделяется слово «кг», сокращенное от «килограмм» (рис. 3). Оно употребляется заметно чаще остальных – 19 раз. Ближайшие по частоте слова – «вес» и «месяц» – употреблены более чем в два раза реже (каждое – по 7 раз). Еще



**Рис. 3. Частотное распределение наиболее популярных слов в проанализированной дискуссии**

реже употребляются слова «неделя» и «год» (4 раза). Другое слово, означающее длительность – «мес», сокращение от «месяц», – употребляется только 3 раза. Оставшиеся слова, представленные на рисунке, употреблены в комментариях дважды. Это слова: «кето», «рост», «считать», «минус», «июнь».

На основании данного набора слов можно заключить, что практика здорового питания тесно связана с практикой похуде-

ния: участники сообщества формируют дискуссию вокруг опыта похудения (слова «кг», «вес»), приобретаемого в течение разных отрезков времени (об этом свидетельствуют словосочетания, описывающие различные промежутки времени).

### *Анализ концептуальных карт*

На рис. 4 представлена концептуальная карта обсуждения в сообществе здорового питания.

Для простоты визуального восприятия представленной коллективной концептуальной карты вершины, не имеющие связей с другими, были предварительно исключены. Так, из 19 исходных комментариев на рисунке представлены только 12. Все связи ориентированы от наиболее ранних к наиболее поздним комментариям, чтобы сохранить причинно-следственную связь и порядок дискуссии. Так, концептуальную карту можно рассматривать как ориентированный несвязный граф, состоящий из трех компонент связности. Вершины связаны 18 ребрами. Общая плотность графа составляет 0,18. Основной темой, раскрываемой в приведенном обсуждении практики здорового питания, является похудение. Данная концептуальная карта фрагментарна. Причем заметно, что обособленные две пары вершин, во-первых, сообщают исключительно факты о количестве сброшенных килограммов за указанный промежуток времени, во-вторых, сильно схожи по структуре.

Перейдем к анализу основного фрагмента обсуждения, в который включается большинство комментариев. Из 12 комментариев, представленных на рис. 4, 8 включены в данный фрагмент и соединены 16 связями. Плотность представленного ориентированного графа составляет 0,57. Данный фрагмент имеет единственную стартовую вершину, которая порождает сразу 5 альтернативных продолжений. Для примера рассмотрим только связи с наибольшим весом со стартовой вершиной цепочки сообщений.



**1. у меня за 6 недель ушло 7,600 кг. решила добавить фото. на первом 2 недели на кето, на втором – сегодня**

- a.** сбросила 23 кг за 3 месяца два года назад. все это время держу вес. три кг туда-сюда гуляет, но это погрешности режима
- b.** за 1 год и 2 месяца 30 кг. фото до и после. было 107 кг, сейчас 77, рост 180 см. я худею очень медленно, сейчас на кето, но уже калории не считаю, считаю только белок и углеводы, жира ем выше нормы. нормы на калькуляторе 149, я под 200 иначе мне голодно. пока прибавки в весе нет, надеюсь так и дальше будет
  - i.** за 2 месяца 10 кг, была 77 сейчас 66.5 рост 179 см, меньше не хочу??
  - ii.** за 2 месяца 3 кг, очень комфортное снижение веса и объёмов. но вес был 61 кг, сейчас 57,5–58 при росте 160.
  - iii.** (сбросила 23 кг за 3 месяца два года назад. все это время держу вес. три кг туда-сюда гуляет, но это погрешности режима)
  - iv.** за полгода 16 кг. уже 3 года вес держится стабильно. ккал не считаю, лень. пью воду. убрала мусорные продукты.
- o** (сбросила 23 кг за 3 месяца два года назад. все это время держу вес. три кг туда-сюда гуляет, но это погрешности режим)

[...]

Автор комментария 1 начинает со стандартной для данного обсуждения структуры: делится количеством килограммов и сроком диеты, добавляя фотографии до и после. Автор комментария а тоже делится успехами похудения, добавляя информацию о прошлом опыте и о нынешнем состоянии. Написанный таким образом комментарий сохраняет структуру противопоставления «до»–«после». Автор комментария b делает то же самое, подробно описывая свои антропометрические данные (рост, а также вес в начале процесса похудения и на момент написания комментария). Также автор подробно описывает строгость соблюдения диеты.

Комментарий b продолжается четырьмя различными сообщениями. При этом комментарии i, ii связаны с предыдущей

вершиной сходным образом: авторы добавляют антропометрическое описание. В отличие от *комментария i*, в *комментарии ii* присутствует субъективная характеристика процесса похудения. В *комментарии iii* связь усиливается, так как в *комментариях b* и *iii* присутствуют описания некоторого пренебрежения диетами. Наконец, *комментарий iv* связан с предыдущим комментарием аналогичным описанием пренебрежения диетами. Порожденная этим комментарием цепочка продолжается рассмотренным ранее комментарием и дополнительно анализироваться не будет.

## *Дискуссия*

Коллективное обсуждение практик здорового питания характеризуется отсутствием выраженного полилога между агентами – они не обсуждают поставленную тему, а, скорее, делятся списком фактов о собственном опыте. Часто этот опыт – некоторые нарушения режима питания, но он не вызывает у агентов никаких особых эмоций. Так, большинство комментариев среди всех проанализированных в концептуальном картировании имеют нейтральную эмоциональную окраску, а тон остальных моделей определить не смогла. Семантически близкие друг к другу сообщения сходны внутренней структурой (срок практикования здорового питания и количество сброшенных килограммов). В дополнение к ней агенты часто упоминают другие антропометрические факты: рост, вес «до» и «после», иногда добавляя фотографии.

## *Заклочение*

Полученные результаты исследования и сделанные на их основе выводы кажутся контринтуитивными. При анализе сообщений здорового питания исследовательская группа ожидала увидеть дискуссии о здоровье и влиянии на здоровье продуктов и пищевых привычек. В действительности же точкой кристалли-

зации дискуссии оказывается похудение. Этот вывод отчетливо прослеживается как в дискурс-анализе, так и в концептуальном картировании. Несмотря на то, что комментарии примерно в равной степени распределены между представителями различных групп агентов, основная дискуссия как борьба дискурсивных позиций развивается между пользователями класса «опытные» и теми, кто только начинает свой путь похудения. Пользователи класса «опытные» чаще остальных участвуют в борьбе за ключевые знаки дискурса. Они же стараются сместить акцент с похудения как процесса изменения веса к его удержанию. Они чаще рассказывают, что вес не теряется, а стоит, килограммы не уходят, а стоят. Этот же вывод оказывается справедлив для представленной в статье концептуальной карты. Здесь важно указать на существующее терминологическое противоречие выбранных методов: агенты могут быть внешне доброжелательны (что может быть обнаружено в результатах концептуального картирования), но имплицитно они будут бороться именно за свою единственно верную интерпретацию того или иного знака. Оттого без явного противоречия с выводами дискурс-анализа мы можем наблюдать, что явное противостояние пользователей не встречается в анализируемой дискуссии. Как правило, пользователи фактологически описывают, каких результатов они достигли в похудении. Они строго связывают здоровье с утратой веса. Вес понимается как нечто ненужное, что следовало бы потерять. Такое понимание веса – как элемента фактологического изложения объемов его утраты – приводится в работе Е. Костяшкиной, где указывается, применительно к медиа-дискурсу, что потеря веса начинает восприниматься как новостной факт. Как и в нашем исследовании, потеря веса в работе Е. Костяшкиной рассматривается в контексте медиадискурса здоровья [19].

В представленной в качестве примера в статье дискуссии существует явный консенсус в понимании здорового питания как процесса, ориентированного на похудение с помощью кетодиеты.

Разногласие происходит в обозначении значимости элементов похудения – что важнее: потерять вес или удержать его? Также разногласия существуют в оценках сложности кетодиеты. Однако концептуальное картирование позволяет получить информацию о принимаемой форме разногласия в конкретной анализируемой дискуссии. Пользователи не вступают в открытый конфликт, а склонны через высказывания о себе популяризировать свою позицию и поддерживать позиции других пользователей, сходные с их собственной. Важно упомянуть, что речь идет о наиболее сильно семантически связанных комментариях. Анализ с включением нерассмотренных комментариев позволил бы получить дополнительные выводы.

В настоящей статье представлено первое приближение пути реализации методов концептуального картирования и дискурс-анализа применительно к анализу онлайн-дискуссий. Опишем подробнее ограничения методов и, соответственно, исследования. Сопоставление результатов дискурс-анализа и концептуального картирования видится небесспорным. С одной стороны, оба метода направлены на единый объект и анализируют один и тот же эмпирический материал – все слова из комментариев концептуальной карты есть на поле дискурса (кроме удаленных стоп-слов). Отсюда методы описывают единую дискуссию о здоровом питании с разных сторон. Однако концептуальное картирование описано только для дискуссии, приводимой в статье в качестве примера. В других дискуссиях в сообществе знаки, выделенные в дискурс-анализе, могут использоваться иначе. Важным ограничением оказывается роль ботов – алгоритмов искусственного интеллекта, автоматически размещающих рекламные и иные сообщения в социальных сетях, нередко выдающих себя за живых пользователей. Эта проблема активно обсуждается в работах С. Брэдшоу и П. Говарда [20], Дж. Прэра [21]. С. Джанвеккио [22] и других авторов. В нашем исследовании это позволяет задать важный вопрос – могут ли эти боты быть субъектами коллективной концептуальной карты,

можем ли мы говорить о разуме таких ботов? Технооптимисты и технопессимисты дают разный ответ на этот вопрос, пишет Д. МакДермотт: у первых искусственный интеллект либо обретет сознание, либо обладает им уже сейчас, у технопессимистов, соответственно, наоборот [23]. Также ограничением выступает выбор авторами уровня анализа в концептуальном картировании: авторы не обобщают высказывания до уровня концептов, что реализовано во многих интерпретациях метода, а остаются на уровне конкретных высказываний. С одной стороны, это позволяет использовать методы поиска семантической близости для обнаружения связи в концептуальной карте и более детально анализировать содержание комментариев, но, с другой стороны, из вида упускается более высокоуровневое обобщение результатов.

Авторы представили в данной статье адаптацию и применение метода дискурс-анализа в интерпретации Лакло и Муфф, а также метода концептуального картирования в интерпретации Трочима для анализа онлайн-дискуссий. Оценка валидности разработанных методов видится задачей для дальнейших исследований.

## ЛИТЕРАТУРА

1. Network analysis reveals open forums and echo chambers in social media discussions of climate change / H.T. Williams, J.R. McMurray, T. Kurz, F.H. Lambert // *Global environmental change*. 2015. № 32. P. 126–138. DOI: 10.1016/j.gloenvcha.2015.03.006.
2. Using social media to monitor mental health discussions – evidence from Twitter / C. McClellan, M. M Ali, R. Mutter [et al.] // *Journal of the American Medical Informatics Association*. 2017. Vol. 24, № 3. P. 496–502. DOI: 10.1093/jamia/ocw133.
3. Trochim W.M., McLinden D. Introduction to a special issue on concept mapping // *Evaluation and program planning*. 2017. № 60. P. 166–175. DOI: 10.1016/j.evalprogplan.2016.10.006.
4. Тумский С.В. Концептуальное картирование как междисциплинарный метод анализа // *Когнитивные исследования языка*. 2014. № 17. С. 182–188. EDN: SALHET.
5. Trochim W.M. An introduction to concept mapping for planning and evaluation // *Evaluation and program planning*. 1989. Vol. 12, №1. P. 1–16. DOI: 10.1016/0149-7189(89)90016-5.

6. *Абрамова Н.В., Николаева Ю.В.* Построение концептуальных карт как метод повышения валидности результатов оценочного исследования // Социология: Методология, методы, математические модели (Социология: 4М). 2006. № 23. С. 83–99. EDN: KVKIXH.
7. *Kane M., Trochim W.M.* Concept mapping for planning and evaluation. CA: Sage Publications, 2007. 216 p. ISBN: 1412940273, 9781412940276.
8. *Lee S., Chun J.* Conceptualizing the impacts of cyberbullying victimization among Korean male adolescents // Children and Youth Services Review. 2020. № 117. art. 105275. DOI: 10.1016/j.chilyouth.2020.105275.
9. *Scourfield P.* A Critical Reflection on the Involvement of “Experts by Experience” in Inspections // The British Journal of Social Work. 2010. Vol. 40, № 6. P. 1890–1907. DOI: 10.1093/bjsw/bcp119.
10. *Йоргенсен М.В., Филлипс Л.Дж.* Дискурс-анализ. Теория и метод / Пер. с англ.; 2-е изд., испр. Харьков: Гуманитарный центр, 2008. 352 с. ISBN: 0-7619-7112-2.
11. *Jones R.H., Chik A., Hafner C.A.* Discourse and digital practices: Doing discourse analysis in the digital age. London: Taylor&Francis, 2015. 262 p. ISBN: 1317537009, 9781317537007.
12. *Carpentier N., De Cleen B.* Bringing discourse theory into media studies: The applicability of discourse theoretical analysis (DTA) for the study of media practises and discourses // Journal of language and politics. 2007. Vol. 6, №. 2. P. 265–293. DOI:10.1075/jlp.6.2.08car.
13. *Игнатьева О.А.* Дискурс-анализ политических суждений в контексте цифровизации // Политическая экспертиза: ПОЛИТЭКС. 2021. Т. 17, № 3. С. 259–272. DOI: 10.21638/spbu23.2021.303. EDN: GBFXQD.
14. *Haditaghi J., Hassasskhah J., Sorahi M.A.* A network-based approach for discourse analysis from Laclau and Mouffe’s perspectives // Journal of Computer-Assisted Linguistic Research. 2020. Vol. 4. P. 1–22. DOI: 10.4995/jclr.2020.12105.
15. *Bakumov P.* An Alternative Model for the Operationalization of Discourse Theory of Laclau and Mouffe // Laboratorium: Russian Review of Social Research. 2022. Vol. 14, №. 3. P. 119–134. DOI: 10.25285/2078-1938-2022-14-3-119-134. EDN: SUGLDK.
16. *Reimers N., Gurevych I.* Sentence-bert: Sentence embeddings using siamese bert-networks // Cornwall University [site]. 27.08.2019. URL: <https://arxiv.org/abs/1908.10084> (date of access: 05.07.2023).
17. *Vatolin A.S., Smirnova E.Y., Shkarin S.S.* Russian News Similarity Detection with SBERT: Pre-training and fine-tuning // Komp’juternaja Lingvistika i Intellektual’nye Tehnologii. 2021. № 20. P. 692–697. DOI: 10.28995/2075-7182-2021-20-692-697. EDN: NKSZTA.
18. GitHub – bureaucratic-labs/dostoevsky: Sentiment analysis library for russian language // GitHub [site]. URL: <https://github.com/bureaucratic-labs/dostoevsky> (date of access: 05.07.2023).

19. Костяшина Е.А. Дискурсивная организация картины мира научно-популярного медицинского журнала // Вестник Томского государственного университета. Филология. 2010. Т. 3, №11. С. 41–47. EDN: NEFMVJ.

20. Bradshaw S., Howard P. Troops, trolls and troublemakers: A global inventory of organized social media manipulation // Computational Propaganda Research Project. 2017. № 12. P. 1–37.

21. Prier J. Commanding the trend: Social media as information warfare. In Information warfare in the age of cyber conflict. London: Routledge, 2020. P. 88–113. ISBN: 9780429470509.

22. Measurement and classification of humans and bots in internet chat / S. Gianvecchio, M. Xie, Z. Wu, H. Wang // USENIX security symposium. 2008. № 17. P. 155–170.

23. McDermott D. Artificial intelligence and consciousness // The Cambridge handbook of consciousness. Cambridge: Cambridge University Press, 2007. P. 117–150. ISBN: 113946406X, 9781139464062.

***Tkach Sergey,***

*Sociologist, Center for Applied Sociology, St Petersburg University,  
St. Petersburg, Russia, s.tkach@spbu.ru*

***Vorobyova Polina D.,***

*Sociologist, Center for Applied Sociology, St Petersburg University,  
St. Petersburg, Russia, st098355@student.spbu.ru*

***Rusakova Maya M.,***

*Candidate of Sociological Sciences, Director, Center for Applied Sociology,  
St Petersburg University, St. Petersburg, Russia, m.rusakova@spbu.ru*

**Experience of implementing discourse analysis and conceptual mapping of healthy eating communities**

The article presents the experience of implementing discourse analysis methods as interpreted by E. Laclau and C. Mouffe and the concept mapping method as interpreted by W. Trochim through the network analysis technique using the example of healthy eating. The result of the analysis is a graph that makes it possible to highlight the struggle of discourse agents for the meaning of key discursive signs, as well as the conceptual map of participants in an online discussion on controversial issues within the topic discussed in the article as an example. The empirical basis for the discourse analysis was 3 000 collected comments in four communities about healthy eating on the social networks VKontakte and Odnoklassniki. The modified versions of the methods of concept mapping and discourse analysis were adapted for the analysis of online discussions. Subsequent validation of the methods seems to be a promising direction for further research. The proposed designs have a number of limitations, which are discussed in the article.

*Keywords:* discourse analysis, concept mapping, graph analysis, sentiment analysis, natural language analysis

**References**

1. Williams H.T., McMurray J.R., Kurz T., Lambert, F. H. Network analysis reveals open forums and echo chambers in social media discussions of climate change, Lambert, *Global environmental change*, 2015, no. 32, p. 126–138. DOI: 10.1016/j.gloenvcha.2015.03.006.
2. McClellan C., Ali M.M., Mutter R., Kroutil L., Landwehr J. Using social media to monitor mental health discussions – evidence from Twitter,

- Journal of the American Medical Informatics Association*, 2017, vol. 24, no. 3, p. 496–502. DOI: 10.1093/jamia/ocw133.
3. Trochim W.M., McLinden D. Introduction to a special issue on concept mapping, *Evaluation and program planning*, 2017, no. 60, p. 166–175. DOI: 10.1016/j.evalprogplan.2016.10.006.
  4. Tumsky S.V. Concept mapping as an interdisciplinary method of analysis (in Russian), *Kognitivnye issledovanija jazyka (Cognitive language research)*, 2014, no. 17, p. 182–188.
  5. Trochim W.M. An introduction to concept mapping for planning and evaluation, *Evaluation and program planning*, 1989, vol. 12, no. 1, p. 1–16. DOI: 10.1016/0149-7189(89)90016-5.
  6. Abramova N.V., Nikolaeva Yu.V. Constructing concept maps as a method for increasing the validity of evaluation research results (in Russian), *Sotsiologiya 4M (Sociology: methodology, methods, mathematical modeling)*, 2006, no. 23, p. 83–99.
  7. Kane M., Trochim W.M. *Concept mapping for planning and evaluation*. CA: Sage Publications, 2007. 216 p. ISBN: 1412940273, 9781412940276.
  8. Lee S., Chun J. Conceptualizing the impacts of cyberbullying victimization among Korean male adolescents, *Children and Youth Services Review*, 2020, no. 117, art. 105275. DOI: 10.1016/j.childyouth.2020.105275.
  9. Scourfield P. A Critical Reflection on the Involvement of “Experts by Experience” in Inspections, *The British Journal of Social Work*, 2010, vol. 40, no. 6, p. 1890–1907. DOI: 10.1093/bjsw/bcp119.
  10. Jorgensen, M. W., Phillips L. *Discourse analysis as theory and method* (in Russian); 2<sup>nd</sup> ed., rev. Kharkov: Gumanitarnyj centr, 2008. 352 p. ISBN: 0-7619-7112-2.
  11. Jones R.H., Chik A., Hafner C.A. *Discourse and digital practices: Doing discourse analysis in the digital age*. London: Taylor&Francis, 2015. 262 p. ISBN: 1317537009, 9781317537007.
  12. Carpentier N., De Cleen B. Bringing discourse theory into media studies: The applicability of discourse theoretical analysis (DTA) for the study of media practises and discourses, *Journal of language and politics*, 2007, vol. 6, no. 2, p. 265–293. DOI:10.1075/jlp.6.2.08car.
  13. Ignatieva O.A. Discourse analysis of political judgments in the context of digitalization (in Russian), *Politicheskaja jekspertiza: POLITJeKS*

- (*Political expertise: POLITEX*), 2021, vol. 17, no. 3, p. 259–272. DOI: 10.21638/spbu23.2021.303.
14. Haditaghi J., Hassasskhah J., Sorahi M.A. A network-based approach for discourse analysis from Laclau and Mouffe’s perspectives, *Journal of Computer-Assisted Linguistic Research*, 2020, vol. 4, p. 1–22. DOI: 10.4995/jclr.2020.12105.
  15. Bakumov P. An Alternative Model for the Operationalization of Discourse Theory of Laclau and Mouffe, *Laboratorium: Russian Review of Social Research*, 2022, vol. 14, no. 3, p. 119–134. DOI: 10.25285/2078-1938-2022-14-3-119-134.
  16. Reimers N., Gurevych I. Sentence-bert: Sentence embeddings using siamese bert-networks, *Cornwall University* [site]. 27.08.2019. URL: <https://arxiv.org/abs/1908.10084> (date of access: 05.07.2023).
  17. Vatolin A.S., Smirnova E.Y., Shkarin S.S. Russian News Similarity Detection with SBERT: Pre-training and fine-tuning (in Russian), *Komp’juternaja Lingvistika i Intellektual’nye Tehnologii*, 2021, no. 20, p. 692–697. DOI: 10.28995/2075-7182-2021-20-692-697. EDN: NKSZTA.
  18. GitHub – bureaucratic-labs/dostoevsky: Sentiment analysis library for Russian language, *GitHub* [site]. URL: <https://github.com/bureaucratic-labs/dostoevsky> (date of access: 05.07.2023).
  19. Kostyashina E.A. Discursive organization of the world picture of a popular scientific medical journal (in Russian), *Vestnik Tomskogo gosudarstvennogo universiteta. Filologija (Bulletin of Tomsk State University). Philology*, 2010, vol. 3, no. 11, p. 41–47.
  20. Bradshaw S., Howard P. Troops, trolls and troublemakers: A global inventory of organized social media manipulation, *Computational Propaganda Research Project*, 2017, no. 12, p. 1–37.
  21. Prier J. *Commanding the trend: Social media as information warfare. In Information warfare in the age of cyber conflict*. London: Routledge, 2020. P. 88–113. ISBN: 9780429470509.
  22. Gianvecchio S., Xie M., Wu Z., Wang, H. Measurement and classification of humans and bots in internet chat, *USENIX security symposium*, 2008, no. 17, p. 155–170.
  23. McDermott D. *Artificial intelligence and consciousness, The Cambridge handbook of consciousness*. Cambridge: Cambridge University Press, 2007. P. 117–150. ISBN: 113946406X, 9781139464062.