

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования

**«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ Н.Г. ЧЕРНЫШЕВСКОГО»**

Кафедра социальной информатики

**ЦИФРОВЫЕ ИННОВАЦИИ И ТРАДИЦИОННЫЕ
ИНСТРУМЕНТЫ ВИЗУАЛЬНОЙ СОЦИОЛОГИИ. АНАЛИЗ
ГРАФФИТИ**

(автореферат бакалаврской работы)

студента 4 курса 451 группы
направления 09.03.03 - Прикладная информатика
профиль Прикладная информатика в социологии
Социологического факультета
Шакурова Дмитрия Сергеевича

Научный руководитель
профессор, доктор социологических наук,
профессор

подпись, дата

О.А. Романовская

Зав. кафедрой
кандидат социологических наук, доцент

подпись, дата

И.Г. Малинский

Саратов 2025

ВВЕДЕНИЕ

Актуальность исследования обоснована стабильно растущим интересом к проблематике информационного общества. Академическое осмысление процессов производства, хранения, переработки и реализации информации, особенно высшей её формы — знаний имеет богатый background. Концептуализация понятия «информационное общество» позволяет непротиворечиво использовать термин «цифровое общество» понимаемое как сетевое информационное общество, отличительной чертой которого является электронно-цифровой способ хранения и распространения информации, электронно-цифровое опосредование большинства социальных отношений.

Одна из основных характеристик цифрового общества заключается в изменении ценностного и экономического фокуса общества с материальных благ на блага нематериальные (информационные). Эта трансформация прослеживается во всех сферах жизни общества: старые практики отбрасываются или видоизменяются, появляются новые.

Вторым важным свойством цифрового общества является увеличение объемов хранимой и производимой информации всех модальностей: текста, аудиальной, визуальной. Активно развиваются направления создания полимодальной информации (мультимодальной информации) типа информации, сочетающей в разных пропорциях классические представления информации, например, аудио и текст, изображения и текст, текст и видео ряд и так далее. Особенно важно отметить увеличение объемов производства визуальной информации. По данным портала Photutorial во всем мире ежедневно делается 5,3 миллиарда фотографий. Доля интернет-трафика, содержащего визуальную информацию или мультимодальную информацию с доминирующей визуальной компонентой так же, увеличивается.

Третьим фактором является парадигматический сдвиг в философии науки, получивший название «визуальный поворот», расширивший значимость визуального как самостоятельного и ценного поля научных исследований, создавая фундамент для изучения разнообразных проявлений визуального.

Важно отметить, что визуальный поворот расширяет границы применимости лингвистических инструментов на пласт более широких явлений, а не полностью смещает фокус исследований. Современная визуальнoцентричная культура не только производит визуальный контент, но и формирует новые социальные практики.

Вместе с возросшим интересом к исследованию визуального перед исследователями возникают новые методологические и методические вызовы. Увеличивающиеся объемы визуальной и мультимодальной информации с визуальным компонентом требуют новых инструментов, позволяющих автоматизировать различные аспекты исследований в области визуальной социологии. Увеличивающиеся из года в год объемы фото и видео материалов не только дают богатую почву для эмпирической базы качественных исследований через анализ открытых источников: соцсетей, личных блогов, сайтов-агрегаторов изображений, но также затрудняют классификацию, отбор и ранжирование материала. Быстро меняющиеся культурные тренды в сфере производства визуального контента требуют инструментов для мониторинга и распознавания подходящих для эмпирической базы случаев. Необходимость проведения процедур поиска, отбора, кодирования и интерпретации при сохранении текущих тенденций и темпов делает затруднительным осуществление этих этапов социологического исследования классическими (ручными) методами или методами частичной автоматизации с привлечением специализированных программных пакетов. Также требуется привлечение большого экономического ресурса.

Общая тенденция, направленная на усиленную цифровизацию общества, дополнительно актуализирует исследовательский интерес к нейросетям с функцией мультимодального ввода. Задачи изучения визуальности и специфика визуальной социологии позволяют рассматривать процессы автоматизации как две подзадачи из области искусственного интеллекта, а именно: задача обработки естественного языка и задачи компьютерного зрения.

Применение разработок из этих двух направлений позволят улучшить такие этапы проведения социологических исследований методами визуальной социологии, как сбор первичной социологической информации, отбор релевантных наборов изображений, соответствующих тематике исследования, учет единиц наблюдения и анализа, кодирование, анализ и обобщение материала. Всё вышеперечисленное позволяет утверждать, о высоком уровне актуальности исследований направленных на развитие методов визуальной социологии в эпоху больших данных

Степень научной разработанности темы: научную разработанность темы можно охарактеризовать как высокую. Визуальная социология имеет богатую методологическую и методическую базы. Классическими работами, заложившими основания для визуальной социологии, можно считать работу Говарда Беккера. Среди методологически значимых работ, осмысляющих различные аспекты визуальной культуры, можно выделить работы Ролана Барта, который рассматривал экзистенциальные и онтологические свойства фотографии как носителя информации. Важным в контексте социального осмысления визуальной информации являются систематическое применение лингвистического инструментария, разработанного в ходе лингвистического поворота. Важный вклад в систематизацию и развитие методологии исследований визуального внес Петр Штомпка, который систематизировал в своих научных публикациях стратегии исследования в рамках интерпретативного и инструментального подходов визуальной социологии.

Среди российских исследователей, разрабатывающих методические и методологические особенности визуальной социологии, можно отметить таких авторов как: Онипко А. А, (изучение сложностей использования визуальных данных как основного источника социологической информации). Н.М. Богданова обращает научное внимание на неконсистентный и слабо формализованный научный аппарат визуальной социологии, однако, успешно применимый в трех основных методологиях социологического исследования.

Ключевым трендом в исследованиях, посвящённых стратегиям организации социологических исследований, является объединение качественно-количественных подходов для формирования целостной и непротиворечивой модели объекта и предмета исследования. Методы визуальной социологии применяются совместно с классическими методами: анкетирование, интервью, фокус-группа. К примеру, в исследовании, посвященном взаимодействию работников и посетителей музея, авторы пишут: «Videography [...] focused on real interaction situations between the guards and the visitors». Это позволяет говорить о применении методов визуальной социологии как необходимости при исследовании сложных контекстно-зависимых систем, а также отделить декларируемое поведение от фактической практики.

Опыты применения нейросетей для решения задач, возникающих в ходе работы с визуальными данными как источником социологической информации, можно охарактеризовать как положительные. Например, исследование «Protest Activity Detection and Perceived Violence Estimation from Social Media Images» активно использует различные детекторы и классификаторы на базе сверточных и рекуррентных нейронных сетей для решения двух задач: обнаружения протестной активности на изображениях из социальных сетей, оценка «насильственности» изображения исходя из видимых признаков. Специализированные нейронные сети выполняющие задачи компьютерного зрения показывают себя успешно при решении таких задач как: распознавание признака и детекция.

Менее изученным является применение проприетарных мультимодальных языковых моделей таких как GPT – 4V для автоматизированной обработки наборов визуальных данных в качественном и количественном подходах. Нами были обнаружены культурологические исследования, направленные на описательный и искусствоведческий анализ предметов искусства. Полученные в ходе экспериментов результаты

свидетельствуют о возможности применения MLLM (Multimodal Large Language Models) для решения описательно-аналитических задач.

Среди работ направленных на выявления границ эвристического потенциала нейросетей с функцией мультимодального ввода (GPT-4V) нами было меньше всего обнаружено исследований в социально-гуманитарном направлении. Таким образом мы обозначаем точку роста для дальнейших исследований, акцентировав внимание на методических и методологических рамках применимости проприетарных нейросетей с функцией мультимодального ввода в качестве инновационного инструмента визуальной социологии.

Объект исследования – методология визуальной социологии в эпоху больших визуальных данных.

Предмет исследования – практика применения нейросети с функцией мультимодального ввода (GPT-4V) в сравнении с классическими методами визуальной социологии при решении исследовательских задач на материалах граффити.

Цель исследования: выявление эвристического потенциала нейросети с функцией мультимодального ввода в сравнении с классическими методами визуальной социологии для решения типовых исследовательских задач.

Для достижения поставленной цели поставлены следующие **задачи**:

- уточнить классические методы визуальной социологии;
- выявить специфику классических методов визуальной социологии в исследовании феномена граффити;
- обосновать технические и теоретические основания для применения нейросетей с функцией мультимодального ввода как инструмента визуальной социологии и провести машинный анализ эмпирического материала;
- провести сравнение результатов применения классического и машинного анализов к эмпирическому материалу – граффити.

Эмпирическая база исследование состоит из:

1. Первичной социологической информации, включающей в себя:

- авторские фотоматериалы, в количестве 158 документов, сделанные в ходе исследования городского пространства Саратова;

- текстовое описание 158 документов, созданных двумя методами: методом автоматической генерации описания изображения на основе распознанного содержания изображения с помощью нейросетей с функцией мультимодального ввода и описания, сделанные «вручную».

2. Вторичной социологической информации, включающей в себя:

- Фотографии граффити из различных журналов и альбомов, посвященных тематике граффити, например выпуски журнала 3rd.Degree.Magazine, Amsterdam, ZGB KAOS.

Структура выпускной квалификационной работы: введение, две главы по три параграфа, заключение, список использованных источников.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

В первой главе «Визуальная социология классический арсенал и его границы.» рассматриваются границы применимости методов визуальной социологии, а также исследуется его положение среди других методов социологии, выявляются основания противоречия, а также предлагаются пути их преодоления.

Так в первом параграфе первой главы нами было проведено критическое осмысление методов визуальной социологии, выявлены основные точки роста на основе библиографического анализа методологических изысканий. Обозначены критически важные положения в эпоху больших визуальных данных. В качестве аргумента, усиливающего изложенную позицию нами, был проведен вторичный анализ статистических данных по созданию фотографий. Вторичный анализ показал устойчивую тенденцию в увеличение объемов производства визуального контента, а также взрывной характер производства.

Обозначена ключевая проблематика, выражающаяся в дисгармоничной динамике между производством визуальных образов и способностью к их познанию и интерпретации. Критически осмыслены стратегии к работе с

визуальной информацией в визуальной социологии в рамках качественного и количественного подходов.

В рамках второго параграфа нами был обоснован выбор феномена, легшего в основу эмпирической базы настоящего исследования. Граффити являются сложным о социокультурно насыщенным артефактом городской среды. Повсеместное распространение феномена делает его доступны и удобным объектом для фиксации методами инструментального подхода. Кросс-культурный и транснациональный характер явления позволяют проводить анализ не только на основе первичных данных, собранных локально, но активно подталкивает к анализу вторичных источников социологической информации: социальных сетей, тематических журналов и фотоальбомов.

Содержательная насыщенность граффити, ярко контрастирующая с повсеместным распространением феномена, а также академические тенденции, направленные на редукционистский подход к исследованию феномена, вызванный сложностью концептуализации и создает испытание не только для мультимодальной языковой модели, но и человека-исследователя.

Отсутствие людей на фотографиях граффити, как основного объекта, содержащего социологическую информацию, не только позволяет применить неклассическую методологию к несвойственному ей объекту, но и рассмотреть роль граффити как практики в рамках поворота к материальному. Отсутствие людей на изображениях, дополнительно повышает надежность результатов инновационного анализа за счет снижения *ai bias*.

В третьем параграфе нами была обозначена проблема сложности концептуализации граффити как объекта социологического исследования. Обозначенная проблематика связана с содержательной насыщенностью феномена, а также междисциплинарным взаимодополняющим характером теоретико-методологических оснований, используемых для анализа и интерпретации феномена.

Нами была предложена трехуровневая схема анализа, состоящая из:

- 1 уровень анализа – объективные, материальные, визуально верифицируемые характеристики феномена.

- 2 уровень анализа – субъективные интерпретационные критерии, сформулированные на основе интерпретативных методик визуальной социологии.

- 3 уровень анализа – синтез и интеретация аналитических единиц, полученных в ходе анализа второго уровня. Теоретической базой третьего уровня анализа становятся теория повседневности Мишеля де Серто и теория пространства Ани Лефевра.

Во второй главе «Эмпирическое исследование: сравнительный анализ результатов применения классических и инновационных методов» речь идёт о применении аналитической схемы, разработанной для интерпретации граффити в третьем параграфе первой главы. Задачей главы является сравнительный анализ результатов применения инновационного (нейросетевого) и классического анализа в рамках одной аналитической схемы.

В первом параграфе проводится классический интерпретационный анализ эмпирического материала, выделяются основные типы граффити, проверяется их соответствие предлагаемой в научной литературе классификации. Выделяются основные образы и наиболее характерные особенности первичного социологического материала.

Во втором параграфе второй главы представлены результаты применения нейросети с функцией мультимодального ввода (GPT-4V) для анализа визуального социологического материала. Модель продемонстрировала высокую эффективность в распознавании визуально верифицируемых признаков граффити — цвета, стиля, техники нанесения, типа поверхности, а также в чтении текстов на изображениях. Эти способности позволяют использовать нейросеть как инструмент первичного анализа и каталогизации визуального материала. Однако при переходе к интерпретации смыслов модель показала ограничения: генерируемые описания зачастую формально убедительны, но лишены глубокого контекстуального основания. Это связано с

различием в природе интерпретации — в то время как исследователь опирается на культурную компетенцию и контекстуальную рефлексия, модель действует на основе вероятностно-статистического сопоставления. Таким образом, установлено, что нейросеть обладает значительным потенциалом в структурной аналитике, но требует дополнения гуманитарной интерпретацией для полноценного социологического анализа.

В третьем параграфе проведён сравнительный анализ результатов, полученных с применением классических и нейросетевых методов визуальной социологии. Установлено, что нейросеть (GPT-4V) демонстрирует высокую эффективность на первом уровне анализа — при фиксации визуально верифицируемых признаков (цвет, техника, материал, шрифт), существенно превосходя ручной анализ по точности и воспроизводимости. Однако при переходе ко второму уровню — интерпретации смыслов — выявлены ограничения: интерпретации, формируемые моделью, часто лишены культурного контекста и глубины. На третьем уровне — теоретического осмысления — нейросеть полностью уступает человеческой рефлексии, будучи неспособной к критическому синтезу. В итоге сделан вывод о комплементарности подходов: нейросетевые инструменты целесообразно использовать на структурно-описательном этапе, в то время как интерпретация и теоретическое обобщение остаются прерогативой исследователя. Предложена гибридная методология, сочетающая сильные стороны автоматизированного анализа с гуманитарной экспертизой.

ЗАКЛЮЧЕНИЕ

В ходе выполнения работы был рассмотрен эвристический потенциал проприетарной нейросети с функцией мультимодального ввода (GPT-4V) для решения типовых задач визуальной социологии на материале граффити. Исследование включало как эмпирическую апробацию алгоритма, так и теоретическое осмысление его применимости в современных методологических условиях.

Были проведены изыскания, направленные на выявление границ применимости визуальных методов в условиях экспансии визуального в культуру и повседневность. Противоречие между масштабами визуального производства и аналитическими возможностями науки стало отправной точкой для размышлений о необходимости гибридных подходов, совмещающих количественную обработку с качественной интерпретацией.

Нейросети, в частности модели типа GPT-4V, открывают возможность к первичной дешифровке визуальных массивов, структурированию признаков, генерации интерпретативных гипотез, которые затем могут быть оценены и осмыслены исследователем. Такая синергия позволяет преодолевать границы классических методов, сохраняя глубину анализа и расширяя охват материала.

В соответствии с поставленной целью были решены следующие исследовательские задачи:

Уточнение классических методов визуальной социологии.

Выполнен теоретический обзор подходов визуальной социологии, акцентирована её методологическая особенность как особого способа «смотреть» на социум сквозь призму визуального. Показана значимость визуальных методов как части более широкого методологического инструментария.

Выявление специфики этих методов в отношении анализа граффити.

Установлено, что граффити представляет собой идеальный эмпирический объект для визуальной социологии: оно богато культурными кодами, контекстуально насыщено, разнообразно по формам и стилям, но при этом устойчиво к формализации и автоматизации анализа. Описаны трудности операционализации.

Обоснование применения нейросетей с мультимодальным вводом.

Предложены теоретические и технические аргументы в пользу использования GPT-4V. Описаны принципы её работы и выявлены функциональные области: дескрипция изображений, распознавание стиля, материалов, контекстов, элементов текста и символов.

Машинный анализ эмпирического материала.

Проведён анализ 158 граффити с помощью нейросети и сопоставлен с результатами ручного анализа. Установлены сильные стороны модели (на уровне распознавания визуальных признаков и генерации описаний), а также её ограничения (на уровне интерпретации).

Сравнение результатов классического и нейросетевого анализа.

Подчёркнуты комплементарность подходов и потенциал гибридной методологии. Выявлено, что наиболее продуктивным является использование нейросети как инструмента первичной обработки и фиксации, за которым следует качественная интерпретация исследователем.