



Б.А. Низомутдинов, А.Б. Углова,
А.С. Тропников

**Автоматизированная обработка
данных для
социально-психологического анализа
цифровой социальной идентичности**

Рекомендуемая форма библиографической ссылки

Низомутдинов Б.А., Углова А.Б., Тропников А.С. Автоматизированная обработка данных для социально-психологического анализа цифровой социальной идентичности // Научный сервис в сети Интернет: труды XXII Всероссийской научной конференции (21-25 сентября 2020 г., онлайн). — М.: ИПМ им. М.В.Келдыша, 2020. — С. 499-514.

<https://doi.org/10.20948/abrau-2020-43>

<https://keldysh.ru/abrau/2020/theses/43.pdf>

Видеозапись выступления

Автоматизированная обработка данных для социально-психологического анализа цифровой социальной идентичности

Б.А. Низомутдинов¹, А.Б. Углова², А.С. Тропников³

1,3 Университет ИТМО, 2 РГПУ им. А.И. Герцена

Аннотация. В статье представлены результаты второго этапа исследовательского проекта, предполагающего разработку методики автоматизированного анализа цифровой социальной идентичности на базе социальной сети «ВКонтакте», а также поиск взаимосвязей между визуальным компонентом социального профиля и психологическими характеристиками респондентов. В ходе исследования осуществлен сравнительный анализ инструментария, основанного на применении технологий машинного обучения, для автоматизированного анализа текстовых и визуальных данных, составляющих основу информационный образ социальной идентичности. С помощью кластерного анализа были выявлены стратегии построения визуальной идентичности. Для выявления цифровых факторов, опосредующих формирование социальной идентичности, был использован корреляционный анализ данных полученных с помощью автоматизированного анализа графических данных и результатов психодиагностического исследования. Стратегии построения визуальной идентичности образовали множественные взаимосвязи с психологическими характеристиками пользователей. На основе полученных взаимосвязей был сделан вывод о широких возможностях автоматизированного анализа цифровой социальной идентичности.

Ключевые слова: автоматизированный анализ графических данных, данных психосоциальная парадигма, социальные сети, социальная идентичность

Automated data processing for socio-psychological analysis of digital social identity

B.A. Nizomutdinov¹, A.B. Uglova², A.S. Tropnikov³

1,3 ITMO University, 2 Herzen University

Abstract. The article presents the results of the second stage of the research project, which involves the development of a methodology for the automated analysis of digital social identity based on the VKontakte social network, as well as the search for the relationship between the visual component of the social profile and the

psychological characteristics of the respondents. In the course of the study, a comparative analysis of the tools based on the use of machine learning technologies was carried out for the automated analysis of text and visual data, which form the basis of the information image of social identity. Using cluster analysis, strategies for building visual identity were identified. To identify digital factors that mediate the formation of social identity, we used a correlation analysis of the data obtained using automated analysis of graphic data and the results of psychodiagnostic research. Strategies for building visual identity have formed multiple relationships with the psychological characteristics of users. Based on the obtained relationships, a conclusion was drawn about the wide possibilities of an automated analysis of digital social identity.

Keywords: parser, automated data processing tools, psychosocial paradigm, social networks, social identity

1. Введение

Современное общество можно охарактеризовать как информационное общество сетевых структур. Интернет представляет собой социальное пространство, в котором рождаются новые формы идентичности. Погружение в Интернет-пространство непосредственно проецируется на личность пользователей. Актуальность исследования состоит в сетевом и виртуальном характере почти всех социальных взаимодействий современного человека, усилении воздействия виртуальных образов на реальную жизнь. Принимая во внимание тенденции к активному вовлечению пользователей в режиме реального времени в пространство социальных сетей, появлению новых стратегий самопрезентации в них, а также возникающих рисках манипуляций с цифровыми образами, возможность автоматизированного анализа социальной идентичности приобретает особую актуальность.

В настоящее время отмечается потребность в новых методах диагностики и прогнозирования различных цифровых трансформаций, затрагивающих формы и механизмы презентации личности в виртуальном пространстве, выявления различных социокультурных рисков, особенностей коммуникативного потенциала социальных сетей. Пространство социальных сетей на данном этапе развития социально-психологических наук является наиболее перспективной площадкой для изучения внутреннего мира пользователей, а также индивидуальных и групповых социальных интеракций. Большинство отечественных исследований, изучающих человека в информационном пространстве опираются на традиционные психодиагностические и экспериментальные методы исследования. Данное исследование направлено на объединение традиционных психодиагностических методов и современных автоматизированных методов обработки и сбора данных для изучения формы и механизмы презентации личности в виртуальном пространстве и дальнейшей возможности прогнозирования социально-психологических характеристик пользователя на

основе анализа его социального профиля без использования трудоемких психодиагностических методик.

Пользователи оставляют в сети большое количество общедоступных текстовых данных, которые могут быть использованы для анализа социально-психологических характеристик, определения уязвимых социальных групп, склонных, например, к антиобщественному, саморазрушающему поведению. Использование современных способов обработки данных, может облегчить подобные исследования, помогая в автоматизированном режиме воссоздавать виртуальную картину мира человека, анализировать предпочитаемые коммуникативные коды пользователя. В предыдущем исследовании нами было проведено пилотажное исследование информационного образа пользователя, которое позволило разработать теоретическую модель автоматизированного анализа данных социальных сетей. Целью данной работы является разработка комплекса автоматизированных средств для анализа факторов формирования цифровой социальной идентичности, на основе анализа текстовой и визуальной информации и соотнесение полученных в автоматизированном режиме данных из социального профиля с результатами психодиагностического экспериментального исследования испытуемых для выявления значимых взаимосвязей их личностных черт с элементами информационного образа.

2. Методология исследования

В настоящее время практически нет междисциплинарных социально-психологических исследований, посвященных автоматизированному анализу структуры цифровых изменений социальной идентичности личности. Имеющиеся исследования закладывают методологическую базу для объединения психологического, социологического и информационного подходов, которые позволят описать цифровую среду как одну из базовых областей жизнедеятельности, задающих направление развития современного человека (Mossberger, Солдатова Г.У.) и ответить на вопрос о степени влияния информационных технологий на психологическое состояние пользователя, его межличностные, групповые и общественные отношения.

2.1. Методологическая модель исследования

Исследование проходило в несколько этапов:

- На первом этапе был осуществлен поиск и анализ инструментария, используемого для решения поставленных задач изучения социальных профилей пользователей, произведен сравнительный анализ автоматизированных систем и тестирование специализированного программного обеспечения.
- На втором этапе были отобраны испытуемые для проведения опроса и экспериментального исследования социально-психологических

характеристик. Было проведено психодиагностическое исследование. Кроме того, респонденты дали свое согласие на анализ и обработку своего профиля в социальной сети Вконтакте.

- На третьем этапе была проведена выгрузка данных из профилей испытуемых, прошедших психодиагностическое исследование на втором этапе. В ходе обработки данных и анализа выходной информации проведено отсеивание неподходящих или неточных сервисов.
- На завершающем этапе, с помощью методов математической статистики, с использованием кластерного и корреляционного анализа были выявлены взаимосвязи между выгруженными данными из социальных профилей и социально-психологическими характеристиками пользователей.

В опросе и психодиагностическом обследовании приняли участие 176 человек, относящихся к первому периоду зрелого возраста (19-32 лет). Средний возраст испытуемых 25,3 года. В исследовании участвовали 58 % (103) женщин и 42 % (73) мужчин.

Был использован традиционный психодиагностический метод для изучения психологических особенностей испытуемых, опосредующий формирование социальной идентичности. Нами был сформирован комплекс психодиагностического инструментария:

- Тест диагностики степени удовлетворенности потребностей Маслоу, который позволил выявить иерархию базовых потребностей индивида, лежащих в основе социальной идентичности;
- Ценностный опросник С. Шварца, направленный на анализ социальных и индивидуальных ценностей, определяющих направленность взаимодействия в сети;
- Self-presentation tactics scale (Lee S.-J., Quigley B., Nesler M., Corbett A., Tedeschi J.), определяющую ведущие стратегии и тактики самопрезентации в информационном пространстве.

В исследовании профиля пользователя в социальной сети были использованы следующие автоматизированные методы:

1. Методы автоматизированного сбора общедоступной информации со страниц профилей в социальной сети ВКонтате. В результате сбора данных был получен массив, который представлял собой статистическую информацию о профиле пользователя социальных сетей. Для сбора информации из социальных сетей использовались 2 подхода:
 - a. Парсинг контента страниц профиля в vk.com/

- b. Методы классификации, основанные на подходе обучения с учителем (Supervised learning) и полуавтоматического обучения с учителем (Semi-supervised learning). Суть задачи заключается в соотношении множества объектов/признаков с заранее заданным множеством классов/категорий. В данном исследовании они были использованы для автоматической классификации сообществ/групп в социальной сети ВКонтакте, в которых состоит пользователь, по различным тематикам более чем в 30 направлениях (Наука, Дизайн, Юмор, Развлечения, Игры, Путешествия и т.д.);
2. Методы обработки графических и текстовых данных для автоматизированного анализа графических данных, направленные на выявление ряда заданных параметров: количество людей на фотографии, эмоции людей, присутствующие объекты, композиция, цветокоррекция, размытие и т.п.

2.2. Методологический анализ инструментария

С позиции инструментария для предложенного исследования имеется определенный массив работ, основанных на применении технологий машинного обучения. Машинное обучение представляет собой широкий раздел области искусственного интеллекта, который изучает методы построения обучающихся алгоритмов. В зависимости от типа сигналов обратной связи или данных, передаваемых системе обучения, выделяют три обширных подхода в машинном обучении (Kumar A., Kingma D.P., Rezende D.J, Mohamed S., Welling M., Ghahramani Z.):

- обучение с учителем (англ. supervised learning) и полуавтоматическое обучение с учителем (англ. semi-supervised learning): машина получает на вход выборку входных данных и ожидаемых результатов, заданных «учителем», цель системы – изучить общие правила приведения входных данных к конечному результату;

- обучение без учителя (англ. unsupervised learning): система пытается самостоятельно найти некую структуру в закономерности данных и сделать некоторый вывод, причем такой подход не предусматривает заранее размеченных наборов данных. Обычно такой подход используется для обнаружения скрытых закономерностей в большом наборе данных;

- обучение с подкреплением (англ. reinforcement learning): программное приложение взаимодействует с изменяющейся окружающей средой, в которой система должна выполнить некоторую определенную задачу без предварительного обозначения конечного результата.

Методами машинного обучения на данный момент решается довольно широкий спектр следующих стандартных задач:

- задача классификации (англ. classification task) (Sokolova M., Lapalme G.) является самым популярным типом задач в машинном обучении и представляет собой соотнесение некоторого множества объектов с метками классов на основе заданной конечной обучающей выборки, заранее разделенной на классы. Различают бинарную и многоклассовую классификации, а также непересекающиеся, пересекающиеся и нечеткие классы;

- задача регрессии (англ. regression task) отличается от задачи классификации тем, что конечным результатом будет действительное число или числовой вектор (промежуток);

- задача ранжирования (англ. learning to rank task) имеет особенность в получении множества выходных данных и автоматической сортировке по значениям ответов; часто используется в информационном поиске и интеллектуальном анализе текстов;

- задача прогнозирования (англ. forecasting task) представляет собой множество объектов, которые являются отрезками временных рядов, заканчивающихся на тот момент, когда необходимо сделать прогноз на будущее значение;

- задача кластеризации (англ. clustering task) (Bakker B., Heskes T.) заключается в разбиении объектов на кластеры (непересекающиеся группы), используя данные о сходстве парных элементов и общие признаки;

- задача фильтрации выбросов (англ. outliers detection task) заключается в обнаружении в обучающей выборке небольшого числа отклоняющихся от нормы объектов; данный подход позволяет избавиться от нежелательных шумов в выборке, неточностей или ошибок в данных;

- задача заполнения пропущенных значений (англ. missing values task) состоит в замене недостающих значений в матрице объектов и их прогнозными значениями.

Машинное обучение как междисциплинарное направление включает в себя такие смежные области знаний, как математическая статистика, методы оптимизации, методы извлечения информации, интеллектуальный анализ данных. Исследование в области машинного обучения проводится путем проведения экспериментов на модельных или реальных данных для проверки правильности работы методов, подтверждения гипотез, получения списка статистически значимых критериев, расчета статистических метрик.

В качестве центральных методов (которые также часто называются подходами или алгоритмами) машинного обучения выделяют следующие (Zhang C., Ma Y., Dietterich T.G.): линейная и логистическая регрессия (linear and logistic regression); метод опорных векторов (SVM, support vector machines); деревья решений (decision trees); случайный лес (random forest); наивный байесовский классификатор (Naive Bayes); градиентный бустинг (boosting); нейронные сети (neural networks); глубокое обучение (deep learning); метод k-

средних (K-means); метод k-ближайших соседей (k-nearest neighbors, KNN); самоорганизующиеся карты (self-organizing maps) и др.

Каждый из представленных методов имеет свои собственные достоинства и недостатки, поэтому они могут применяться для решения совершенно разных типов задач. Обработка естественного языка (англ. Natural Language Processing, сокращенно NLP) используется в сочетании с методами машинного обучения для выполнения таких задач, как обнаружение эмоций, анализ тональности текста, распознавание актов диалога, классификация спама в письмах, машинный перевод, распознавание речи и т. д. NLP играет очень важную роль в сборе, обработке и анализе данных путем преобразования естественного языка в формат, который в дальнейшем используется методами машинного обучения для реализации собственных алгоритмов. Так, Йоав Голдберг (Yoav Goldberg) в своих работах, а также Т. В. Батура в обзорном исследовании методов автоматической классификации текстов и А. О. Зиберт с В. И. Хрустальевым приводят основные методы и подходы обработки естественного языка (NLP) в рамках работы с нейронными сетями разных архитектур и со стандартными статистическими моделями для реализации методов глубокого обучения, а также приводят основные результаты, полученные в ходе реализации данных методов.

Среди таких методов и подходов можно выделить следующие: токенизация; составление списка стоп-слов; стемминг; лемматизация; извлечение именованных сущностей (англ. Named Entity Recognition); модель «мешок слов»; вычисление функции TF-IDF; алгоритмы Word2Vec; др.

Основываясь на ряде проводимых исследований в области выявления взаимосвязей социальных сетей и личностных характеристик пользователей (Kosinski M., Zhang C., Settanni M., Azucar D., Marengo D.) можно выделить несколько направлений по определению взаимосвязей данных, полученных из социальных сетей и персональными качествами пользователей:

- обработка фотоизображений
- семантический анализ текстовых пользовательских «постов»
- анализ статистических данных.

Обработка данных таких масштабов требует значительных трудозатрат. Для автоматизации данного процесса, и ускорения всего процесса определения искомых корреляционных связей, целесообразно применение облачных платформ, позволяющих удаленно направлять запросы на проведение комплексных операций по анализу данных. Подобные платформы по анализу графических изображений предоставляют Microsoft, Google, Amazon. Microsoft Azure – открытая платформа, предоставляющая услуги для облачных вычислений и специализированных сервисов, разработанных Microsoft.

Microsoft Azure Cognitive Services – интеллектуальные интерфейсы API, специализирующиеся на машинном анализе, основной функцией которых является обработка поступающих данных и выявления различной информации.

Облачные службы «Распознавание лиц» и «Компьютерное зрение» используются для обработки фотографий и выявления ряда данных.

Уже приводящиеся исследования по использованию Microsoft Azure Cognitive Services с целью выявления личностных качеств человека выяснили с помощью облачного сервиса что существует взаимосвязь между количеством лиц на изображениях в профиле социальных сетей и чертами экстраверта. Так же была установлена корреляция между набором технических данных изображений (размытие, резкость) и возрастом, полом и уверенностью пользователя. Microsoft Azure Cognitive Services предоставляет для бесплатного пользования пакет на 20 транзакций в минуту, общим количеством 5 000 и 50 000 транзакций в месяц для сервиса «Компьютерное зрение» и «Распознавание лиц», соответственно. Существует также пакет для студенческого пользования с расширенным функционалом. Подключение осуществляется через API.

В 2008 году компания Google представила свою площадку по предоставлению услуг облачных вычислительных мощностей и сервисов под названием Google Cloud Platform. Являясь прямым конкурентом облачной площадки Microsoft Azure, площадка также предоставляет услуги хостинга, обработки массивов данных и т.д. Службой, отвечающей за обработку и анализ изображений, является служба Google Cloud Vision.

Как и аналоги подобной службы, Cloud Vision способен определять лица на изображениях, анализировать проявляемые эмоции, фиксировать отличительные черты в виде волосяного покрова или головного убора, выявлять предметы и обстановку на заднем фоне, а также составлять набор «тэгов» для изображения. В отличие от Microsoft Azure, Google Cloud Vision специализируется на анализе существующих объектов при обработке изображении, а не на простой фиксации. Таким образом, набор конечных тэгов для двух сервисов различается: сервис Microsoft перечисляет все возможные объекты на изображении и выводит краткое описание происходящего, в то время как сервис Google выводит уже проанализированный набор тэгов, полученный из определенных объектов. Amazon Web Services является площадкой по предоставлению услуг машинной обработки данных от компании Amazon. Данная платформа специализируется на предоставлении услуг по организации облачных баз данных, бессерверных вычислений, средств разработки, виртуальных серверов и хранилищ. На данный момент Amazon Web Services сконцентрирован на предоставлении инфраструктурных и платформенных услуг. Тем не менее, AWS предоставляет технологию Amazon Rekognition, предназначенную для интеллектуального анализа изображений и видео.

Amazon Rekognition предоставляет стандартный функционал сервиса машинной обработки изображений. С его помощью можно определять лица, предметы, тэги, степень пристойности изображения, цветовую гамму и ряд других технических характеристик.

В ходе проведения экспериментального тестирования и изучения аналитических статей консалтинговых компаний, было установлено что данный сервис уступает описанным двум предыдущим в точности определения отображаемых объектов.

Amazon Rekognition способен лишь фиксировать объекты на изображении, но не анализировать общую картину, в отличии от Google Cloud Vision. Также, технология от Amazon не способна составлять смысловое описание изображения как конкурентная служба от Microsoft. При этом, общая точность зафиксированных тэгов относительно низкая.

Основываясь на ряде проведенных исследований (Sophie W. F., Xenos S., Ryan T.) можно установить, что лингвистические особенности могут быть использованы для распознавания личностных характеристик. Описанные методы могут быть использованы не только при анализе рукописных или машинописных текстов, но также и текстов, оставляемых пользователями социальных сетей.

В основном большинстве проводимых исследований по этой тематике используется программа “Linguistic Inquiry And Word Count”, позволяющая посчитать доли определенных частей речи в тексте, количество слов длинее определенного значения, количество используемых знаков препинания, количество слов из различных лексико-семантических категорий. Далее, полученные данные соотносятся с данными, полученными в ходе тестирования авторов текста и в ходе обработки выделяется корреляция между определенными лингвистическими особенностями и личностными характеристиками.

Основной проблемой является отсутствие общедоступных и проработанных библиотек русского языка для подобного рода программ. Тем не менее, используя схожие по функционалу онлайн-сервисы, можно с меньшими трудозатратами обрабатывать большее количество текстовой информации.

Microsoft Azure, описываемый ранее в качестве платформы по предоставлению облачных вычислительных услуг, также имеет службу по анализу текстовых сообщений. Полученные данные можно использовать для определения отношения человека к тем или иным объектам или событиям, тематику его текстовых сообщений, круг интересов и общий тон публикаций.

ISPRAS API – это некоммерческий продукт, разработанный Институтом системного программирования им. В.П. Иванникова РАН. Данный продукт специализируется на обработке и анализе естественного языка. На текущий момент, существует несколько демоверсий различных программных решений, одна из которых направлена на проведение семантического анализа текста – Text Processing. Данный программный продукт реализован на некоммерческой основе и предлагает более внушительный функционал, чем аналоги. Получаемые в ходе обработки текстовых сообщений данные можно будет использовать не только при определении круга увлечений пользователя или его

отношения к тем или иным объектам, но также и при поиске взаимосвязей между частотой употребления определенных частей речи и личностными чертами характера.

Одной из наиболее известных компаний, занимавшихся машинным анализом пользовательских данных являлась Cambridge Analytica (CA), созданная на основе исследований Михала Косински. Косински проводил исследования с помощью приложений в социальных сетях, предлагая пользователям проходить различные психологические тесты. Собирая с разрешения пользователей их данные, он смог доказать, что между деятельностью человека в сети и его реальным "альтер-эго" существует связь. Так, например, с помощью 68 "лайков" можно с определенной долей вероятности установить пол, возраст, сексуальную ориентацию и политические предпочтения пользователя. На основе данных исследований была разработана система, применявшаяся CA. Используя данные пользователей из социальных сетей и основываясь на результатах исследования Михала Косински, специалисты CA смогли спрогнозировать их информационные образы. Далее, информационные образы сегментировались по различным параметрам (политическим предпочтениям, психологическим чертам и т.п.), после чего данные об отсортированных пользователях продавались третьим компаниям - для демонстрации таргетинговой политической рекламы, которая была создана специально под отдельные категории людей.

Используя этот опыт, многие Российские компании начали предлагать свои услуги по идентификации и сегментации пользовательского контингента. Так, глава Сбербанка заявил, что компания намерена применять методы Михала Косински для идентификации клиентов с высоким риском просрочки выплат.

На данный момент, на Российском рынке можно выделить компанию Double Data, которая занимается сбором данных из социальных сетей для снижения уровней просроченных задолженностей и выявления мошенничества в розничном кредитовании за счет использования новейших технологий и встраивания их в существующие банковские бизнес-процессы.

3. Результаты исследования

На первом этапе автоматизированного анализа профилей социальных сетей был проанализирован визуальный контент пользователей с помощью платформы Microsoft Azure Cognitive Services, который позволил выделить 227 тегов, встречающихся на аватарах пользователей. Нами были отобраны основные 46 тегов (объекты), которые встречаются у большинства испытуемых. Набор данных тегов представляет собой базовую визуальную идентичность пользователей ВКонтакте, которая может стать уникальным источником информации о социальной жизни человека, его гендерных, возрастных, политических особенностях для различного рода социально-психологических исследований.

Для выявления основных стратегий построения визуальной идентичности, представленных на фотографиях, объединяющих, полученные в ходе автоматического анализа теги, использовался кластерный анализ (метод Варда) (рисунок 1).

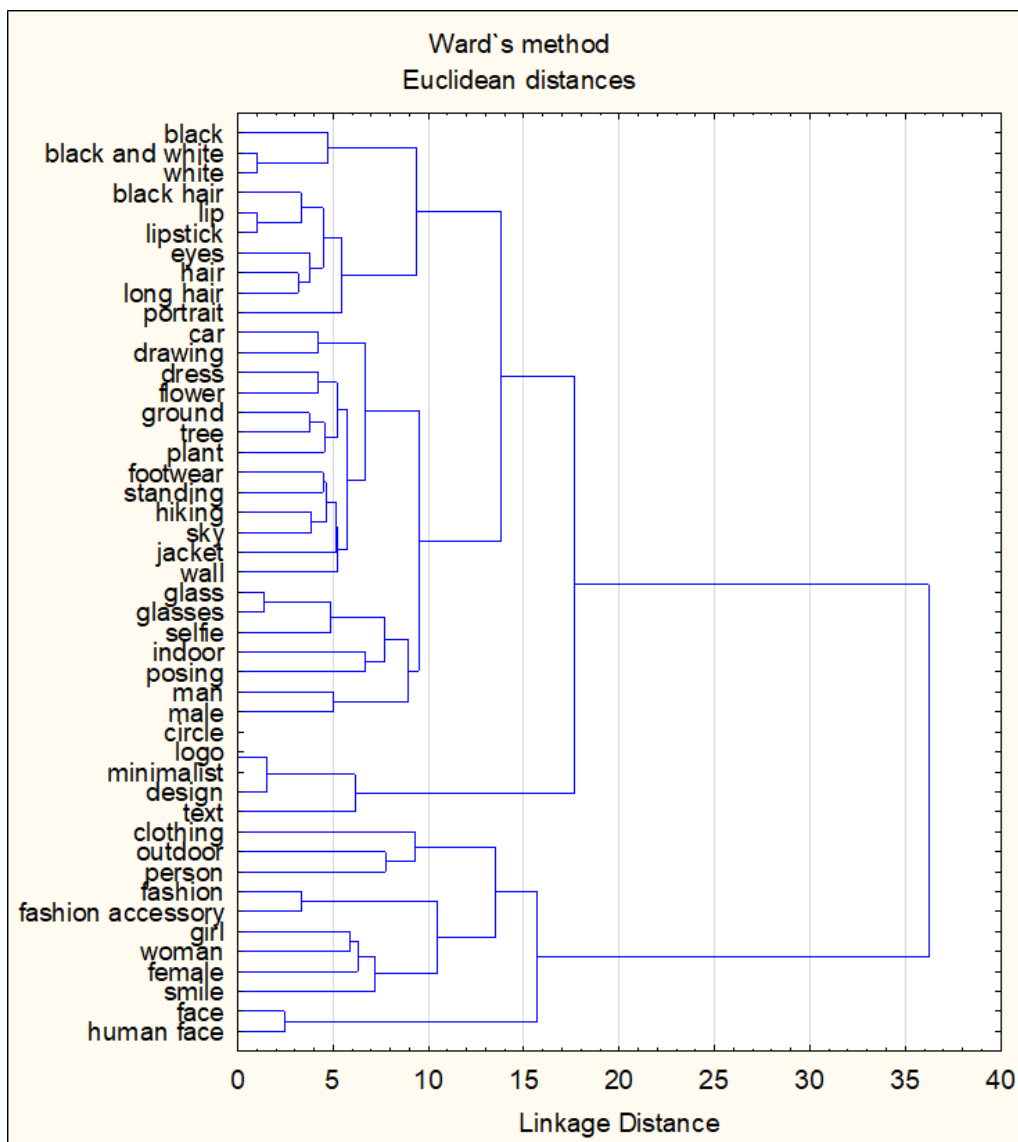


Рисунок 1. Категориальная структура визуальной идентичности пользователей ВКонтакте

Кластерный анализ позволил выделить и описать основные стратегии построения визуальной идентичности в социальной сети:

- Деловая, портретная самопрезентация (black, black and white, black hair, white, lip, lipstick, eyes, long hair, portrait, hair), призванная создать определенную дистанцию между пользователем и зрителем;
- Визуальное повествование (car, drawing, dress, flower, ground, plant, tree, footwear, standing, hiking, sky, jacket, wall), призванное за счет природного антуража донести эмоционально-образное сообщение, понравиться зрителю;

- Постановочная женская самопрезентация (clothing, face, fashion, fashion accessory, girl, human face, outdoor, person, smile, woman, female), является отражением регламентированной социально-ориентированной идентичности и презентует набор популярных клише, для создания нормативного образа;
- Постановочная мужская самопрезентация (glass, glasses, selfie, indoor, posing, man, male), также является отражением регламентированной социально-ориентированной идентичности;
- Анонимная самопрезентация (design, logo, minimalist, text), призванная указать на нежелание пользователя расширять круг социальных контактов, ориентированного на узкий круг пользователей.

На следующем этапе с помощью корреляционного анализа нами были выявлены достоверно значимые взаимосвязи между выявленными стратегиями построения визуальной идентичности и элементами реальной социальной идентичности, изученной в ходе психодиагностического исследования:

1. Деловая, портретная самопрезентация образовала достоверно значимые взаимосвязи с такими социально-психологическими характеристиками как «количество времени, проводимое в социальных сетях» ($r=0,16$, $p \leq 0,05$), «потребность в самоактуализации» ($r=0,22$, $p \leq 0,05$).
2. Стратегия «Визуальное повествование» образовала множественные взаимосвязи с такими социально-психологическими характеристиками как «Соответствие профиля в социальной сети реальному Я» ($r=0,17$, $p \leq 0,05$), «ценность самоуважения» ($r=0,17$, $p \leq 0,05$), «ценность единства с природой» ($r=0,18$, $p \leq 0,05$).
3. Постановочная женская самопрезентация образовала положительные достоверно значимые взаимосвязи с такими социально-психологическими характеристиками как «количество используемых социальных сетей» ($r=0,19$, $p \leq 0,05$), «ценность поиска смысла жизни» ($r=0,17$, $p \leq 0,05$), «ценность самоуважения» ($r=0,26$, $p \leq 0,05$), «ценность единства с природой» ($r=0,16$, $p \leq 0,05$), «ценность принятия жизни» ($r=0,16$, $p \leq 0,05$), «ценность честности» ($r=0,19$, $p \leq 0,05$), «потребность в самоутверждении» ($r=0,19$, $p \leq 0,05$), «потребность в самоактуализации» ($r=0,17$, $p \leq 0,05$), «самоценность» ($r=0,19$, $p \leq 0,05$), «личностная зрелость» ($r=0,18$, $p \leq 0,05$).
4. Постановочная мужская самопрезентация образовала отрицательные достоверно значимые взаимосвязи с такими социально-психологическими характеристиками как «ценность чувства принадлежности» ($r=-0,17$, $p \leq 0,05$), «ценность уважения мнения других» ($r=-0,20$, $p \leq 0,05$), «ценность скромности» ($r=-0,22$, $p \leq 0,05$) и положительную взаимосвязь с «тактикой взаимодействия – запугивание» ($r=0,16$, $p \leq 0,05$).

5. Анонимная самопрезентация образовала отрицательные достоверно значимые взаимосвязи с такими социально-психологическими характеристиками как «количество используемых социальных сетей» ($r=-0,17$, $p \leq 0,05$), «ценность самоуважения» ($r=-0,22$, $p \leq 0,05$), «ценность послушания» ($r=-0,19$, $p \leq 0,05$), «самопринятие» ($r=-0,18$, $p \leq 0,05$), а также положительную взаимосвязь с «ценностью свободы» ($r=-0,16$, $p \leq 0,05$)

В результате можно говорить о том, что были выявлены многочисленные достоверно значимые взаимосвязи между компонентами социальной идентичности, выявленными при помощи психодиагностических методик и стратегиями построения визуальной идентичности, выявленными с помощью автоматизированного анализа страниц пользователей, что может в дальнейшем облегчить проведение социально-психологических исследований.

4. Заключение

Значимость данного исследования состоит в том, что его результаты могут поспособствовать разработке и апробации инструментария для междисциплинарных научных исследований, позволяющего строить прогностические модели поведения в сети Интернет, опосредующие формирование социальной идентичности.

Исследовательский интерес на последующих этапах работы следует сконцентрировать на качественном анализе собранных данных, а также выявлении поиска неочевидных закономерностей и корреляций в исходных данных с помощью использования методов машинного обучения для ускорения получения и обработки информации о социальной идентичности в информационном пространстве.

Литература

1. Войскунский А., Евдокименко А., Федунина Н. Сетевая и реальная идентичность: сравнительное исследование // Психология. Журнал Высшей школы экономики. — 2013. — Т. 10, № 2. — С. 98–121
2. Емелин Вадим Анатольевич Высшие психические функции в контексте цифровых // Цифровое общество в культурно-исторической парадигме, коллективная монография. Под редакцией Т.Д. Марцинковской, В.Р. Орестовой, О.В. Гавриченко. Москва, 2019, с. 177-181
3. Летов Е.В. Сетевая идентичность в контексте культурных процессов информационного общества: автореферат дисс... канд. философ. наук. — Москва, - 2013. — 19 с.
4. Низомутдинов Б.А., Тропников А.С., Углова А.Б. Разработка прогностической модели информационного образа пользователя с

- применением автоматизированных средств обработки данных из социальных сетей // Научный сервис в сети Интернет. 2019. № 21. С. 532-540.
5. Соловьева Л.Н. Цифровая идентичность как новый вид идентичности человека информационной эпохи // Общество: философия, история, культура. 2018 doi: 10.24158/fik.2018.12.6
 6. Углова А.Б., Королева Н.Н., Богдановская И.М., Луговая В.Ф. Стратегии виртуальной самопрезентации современных российских учителей в социальных сетях // Письма в Эмиссия.Оффлайн: электронный научный журнал. 2019. № 10. С. 2773
 7. Cote J. E., Schwartz S. J. Comparing psychological and sociological approaches to identity: Identity status, identity capital, and the individualization process // Journal of Adolescence. 2002. № 25. P. 571-586.
 8. Katz, J. E. & Rice, R. E. Social consequences of Internet use: Access, involvement and interaction. Cambridge, MA: The MIT Press, 2002.
 9. Kaiqi H., Qiao W., Zhenyang W. Natural color image enhancement and evaluation algorithm based on human visual system // Computer Vision and Image Understanding. - V. I, N 103. - 2006. - P. 52-63.
 10. Pianesi F., Mana N., Cappelletti A., Lepri B., Zancanaro M. Multimodal recognition of personality traits in social interactions // Proc of the 10th international conference on Multimodal interfaces. - 2008. - P. 53-60.
 11. Settanni M., Azucar D., Marengo D. Predicting the Big 5 personality traits from digital footprints on social media: A meta-analysis // Personality and Individual Differences. - 2017. - P. 150-159.
 12. Sophie W.F., Susanne B.E., Jochen , Patti V. Norms of online expressions of emotion: Comparing Facebook, Twitter, Instagram, and WhatsApp // SAGE. - N 20. - 2017. - P. 1-19.
 13. Tropnikov A., Uglova A., Nizomutdinov B. Development of a prognostic model of the user's information image using automated tools for processing data from social networks // Communications in Computer and Information Science. 2019. T. 1038. С. 405-413.
 14. Vartanova E.L. The media and the individual: economic and psychological interrelations // Psychology in Russia: State of the Art. - 2013. - 6(1). - 110-118.
 15. Vapnik V.N. An Overview of Statistical Learning Theory // Neural Networks, IEEE Transactions on. 1999. Vol. 10. № 5. p. 988–999
 16. Voiskounsky A.Ye. (2013). Psychology of computerization as a step towards the development of cyberpsychology // Psychology in Russia: State of the Art. 2013. - 6(4). - 150-159.

17. Xenos S., Ryan T. Who uses Facebook? An investigation into the relationship between the Big Five, shyness, narcissism, loneliness, and Facebook usage // *Computers in Human Behavior*. - V. 27, N 5. - 2011. - P. 1658-1664.
18. Wexner L. The degree to which colors (hues) are associated with mood-tones. // *Journal of Applied Psychology*. - V. VIII, N 38. - 1954. - P. 432-435.