

Федеральное государственное бюджетное учреждение науки
Центральная научная библиотека
Уральского отделения Российской академии наук

**IX Информационная школа
молодого ученого**

Сборник научных трудов

Екатеринбург
2021

DOI 10.32460/ishmu-2021-9-0000
ББК 72.471.1(235.55)я431(2)
УДК 001(082)
И 74



И 74 **Информационная школа молодого ученого (9, 2021; Екатеринбург)**
IX Информационная школа молодого ученого : сб. науч. тр. / отв. ред.
П. П. Трескова ; сост. Ю. Д. Прокофьева, Т. А. Осипенко, Л. А. Оболенская,
О. Н. Зырянова, М. А. Пекшева, В. Ю. Герасименко, А. Ю. Герасименко,
Н. А. Парменова. Екатеринбург, 2021. 519 с.

Редакционная коллегия:

П. П. Трескова, кандидат педагогических наук (отв. ред.)
Т. И. Горбунова, доктор химических наук
А. Г. Кислов, кандидат философских наук
А. С. Луньков, кандидат исторических наук
А. А. Старостин, кандидат физико-математических наук
М. В. Модоров, кандидат биологических наук
Н. Б. Граматчикова, кандидат филологических наук
М. Н. Макарова, кандидат экономических наук
Д. В. Волосников, кандидат физико-математических наук
М. А. Болков, кандидат медицинских наук
Е. К. Созина, доктор филологических наук

Рецензент

С. И. Кумков, кандидат технических наук

ISBN 978-5-8295-0788-6

Сборник научных трудов включает материалы, представленные на Всероссийской междисциплинарной молодежной научной конференции с международным участием «IX Информационная школа молодого ученого» (Екатеринбург, 20–23 сентября 2021 г.). Конференция организована Центральной научной библиотекой Уральского отделения Российской академии наук и Секцией специальных научных, научно-технических и технических библиотек Российской библиотечной ассоциации в рамках Года науки и технологий в Российской Федерации. Материалы сборника содержат результаты исследований молодых учёных в различных отраслях естественных, технических и гуманитарных наук.



ББК 72.471.1(235.55)я431(2)
УДК 001(082)

ISBN 978-5-8295-0788-6

© ЦНБ УрО РАН, 2021

© Коллектив авторов, 2021

DOI 10.32460/ishmu-2021-9-0055

УДК 314.02

ББК 60.72

Источники демографических данных в условиях цифровизации экономики и общества*

*Смирнов А. В., канд. экон. наук, ст. науч. сотр.,
ИСЭ и ЭПС ФИЦ Коми НЦ УрО РАН, Сыктывкар*

Ключевые слова: демография, источники данных, цифровизация, большие данные, миграция, Арктика

Рассматриваются новые источники демографических данных, возникшие в результате цифровой трансформации экономики и общества. Систематизированы виды интернет-сервисов, применяемых в социально-демографических исследованиях. Выявлены возможности и ограничения их использования в качестве источников демографических данных. Рассмотрен пример изучения межмуниципальной миграции населения российской Арктики по данным социальных медиа.

Sources of demographic data in the context of digitalization of the economy and society

*Smirnov A. V., PhD in Economics, senior researcher,
ISE&EPN FRC Komi SC UB RAS, Syktyvkar*

Keywords: demography, data sources, digitalization, big data, migration, Arctic

The article considers new sources of the demographic data that have emerged as a result of the digital transformation of the economy and society. The article systematizes types of the Internet services used in the socio-demographic research. The study reveals possibilities and limitations of their use as sources of the demographic data. The author gives an example of studying the intermunicipal migration of the Russian Arctic population according to the social media data.

* Исследование выполнено при поддержке гранта Российского научного фонда № 21-78-00081.

Введение

К 2019 году доля населения планеты, использующего интернет, превысила 50%, а в развитых странах приближается к 90%¹. С помощью цифровых технологий человек осуществляет все больше повседневных задач, таких как получение новостей, общение, развлечение, учеба, поиск работы [1]. В процессе использования интернет-сервисов люди оставляют в сети множество данных, которые могут использоваться в исследовательских целях. В демографии, как и в других общественных науках, в настоящее время большое внимание уделяется осмыслению происходящих изменений, разработке инструментария, который позволит применять новые цифровые источники данных в изучении демографических процессов.

В данной статье систематизированы источники демографических данных, возникшие в результате цифровой трансформации экономики и общественной жизни. Показаны их сильные и слабые стороны, перспективы применения. Кроме того, рассмотрен пример изучения маршрутов миграционных перемещений на уровне отдельных муниципальных образований, что недоступно с помощью данных официальной статистики.

Цифровые источники демографических данных

Источники демографических данных в цифровой среде можно разделить на две большие группы. Во-первых, это традиционные демографические данные (итоги переписей, официальная статистика, обследования населения), размещенные в цифровом виде. Эти данные создавались непосредственно в исследовательских целях. Цифровая среда позволяет в удобном для обработки виде размещать не только итоги наблюдений, но и микроданные, что открывает для исследователей новые возможности. В России примерами могут служить базы микроданных переписей населения, публикуемые Росстатом, а также Российский мониторинг экономического положения и здоровья населения НИУ ВШЭ, включающий лонгитюдные данные, начиная с 1994 года. Доступ к микроданным позволяет изучать распределения признаков любого вида, не ограничиваясь заранее определенным набором шаблонов.

Ко второй группе данных, принципиально новым источникам, полезным в демографических исследованиях, можно отнести различные интернет-сервисы (таблица 1). При использовании интернет-сервисов

¹ По данным Международного союза электросвязи. URL: [https://www.itu.int/en/ITU-D/Statistics/Pages/stat\(дата обращения: 20.06.2021\)](https://www.itu.int/en/ITU-D/Statistics/Pages/stat(дата%20обращения:20.06.2021)).

пользователями генерируется цифровой след, который «предоставляет беспрецедентные возможности для сбора как экспериментальных, так и наблюдательных данных в масштабах, одновременно огромных и микроскопических» [2, с. 131]. Учитывая большой масштаб накапливаемой информации, зачастую исследователям приходится прямо или косвенно работать с «большими данными».

Таблица 1

**Интернет-сервисы, применяемые
в демографических исследованиях**

Виды сервисов	Примеры сервисов	Направления исследований
Анализаторы поисковых запросов	Яндекс. Подбор слов, GoogleTrends	Миграционные и репродуктивные установки населения по территориям на основе запросов пользователей поисковых систем
Социальные сети	ВКонтакте, Facebook, Twitter, Instagram	Изучение демографических установок, топологии сетей взаимодействий населения, новых видов мобильности. Изучение миграций на основе изменений геолокации и анкет пользователей.
Картографические сервисы	Яндекс.Карты, Google Earth, NightEarth.com, NASA Earth Observatory	Анализ изменений пространственного размещения населения и уровня жизни по авиа- и спутниковым снимкам, в том числе ночным.
Тематические сайты	Интерактивные порталы служб занятости, LinkedIn	Изучение миграционных и репродуктивных планов на основе статистики спроса и предложения сайтов вакансий, продажи/аренды недвижимости и др.
Сайты с генеалогическими данными	WikiTree	Изучение влияния родственных связей на рождаемость, продолжительность жизни и пространственную мобильность.

Виртуальное население, данные о котором накапливаются сервисами, не соответствует реальному, а представляет собой его проекцию, искаженную цифровой средой. С этим связаны многие недостатки цифровых источников данных, среди которых низкая репрезентативность, фрагментарность, уязвимость к изменениям, ошибки в работе алгоритмов, ложные сведения, спам-аккаунты, низкая достоверность, дублирующие данные, ограниченность доступа к данным [2, 3]. Сильные

стороны «больших данных» с точки зрения общественных наук обобщены Р. Китчиным. Это огромный объем, высокая скорость производства, высокая детализация, внутреннее многообразие, исчерпывающая полнота, взаимосвязь с другими данными, расширяемость, масштабируемость [4]. Кроме того, новые источники данных лишены одного из главных ограничений классических опросов – реактивности – респонденты знают, что участвуют в исследовании, что может оказать эффект на результаты [5].

Данные могут обрабатываться как традиционными демографическими методами и моделями, так и с применением инструментария «науки о данных». Это прежде всего интеллектуальный анализ данных, машинное обучение, алгоритмическое и вероятностное моделирование. В рамках одного исследования может использоваться множество методов. Разрабатываются методы верификации цифровых демографических данных. Например, сопоставление с официальной статистикой позволяет устранить расхождения путем учета распространенности цифровой платформы в различных социальных группах или регионах. Выборка разбивается на классы и взвешивается в соответствии с известными демографическими показателями (пост-стратификация). В отсутствие официальной статистики, если структура выборки меняется относительно медленно, можно ограничиться тенденциями изменения показателей во времени. При наличии множества источников информации может применяться байесовское моделирование [6].

Методика исследования

Официальная статистика о направлениях миграционных перемещений в разрезе территорий прибытия и выбытия публикуется Росстатом только на уровне федеральных округов и субъектов Российской Федерации, но не на уровне муниципальных образований, что затрудняет проведение детального анализа межмуниципальных миграций. Новые цифровые источники данных способны частично восполнить этот пробел. Воспользуемся данными проекта «Виртуальное население России»², реализованного при поддержке Русского географического общества. Путем обработки анкет пользователей самой популярной в России (на момент сбора данных) социальной сети «ВКонтакте», получены данные о миграции населения за январь-март 2015 г. в привязке к городским округам и муниципальным районам. Около 88 млн анкет содержали данные о месте жительства (или последнем месте учебы,

² Виртуальное население России. URL: <http://webcensus.ru/> (дата обращения: 20.06.2021).

если место жительства не было указано), расположенном на территории Российской Федерации [7].

С помощью данных проекта «Виртуальное население России» рассмотрим маршруты миграции населения российской Арктики, включающей 75 городских округов и муниципальных районов в девяти субъектах РФ. Арктические территории характеризуются крайне высокой миграционной подвижностью при сокращении численности постоянного населения [8; 9]. Алгоритм расчета показателей и визуализация графа миграций реализованы автором на языке программирования Julia с использованием пакетов DataFrames.jl, LightGraphs.jl и GraphPlot.jl.

Анализ межмуниципальных миграций

Стрелки на рисунке 1 показывают наиболее популярные направления перемещений из каждого муниципального района или городского округа в российской Арктике. Жители 23 муниципальных образований (31%) чаще всего перемещаются в Москву и Санкт-Петербург. Это в основном большие городские округа, жители которых имеют ресурсы и возможности для переезда в федеральные центры. Причем Санкт-Петербург более

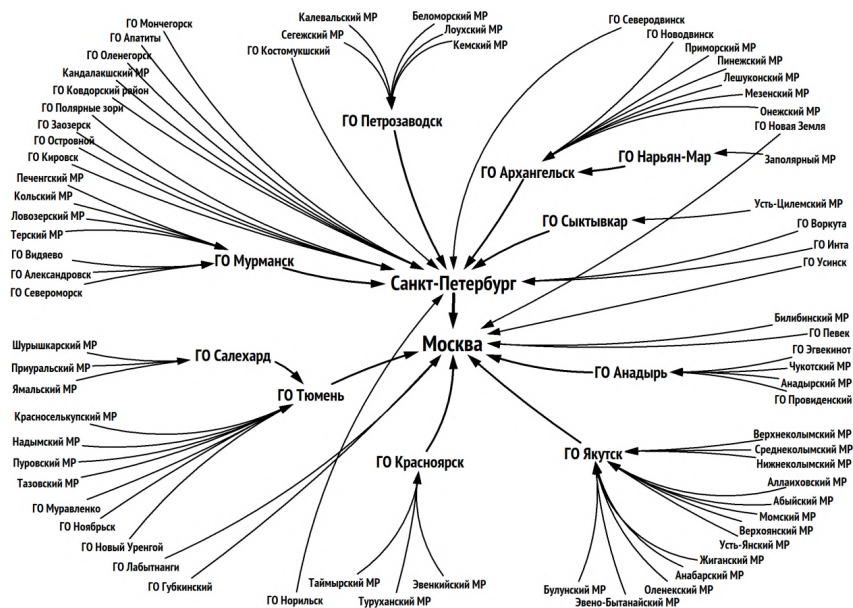


Рис. 1. Основные маршруты межмуниципальной миграции в российской Арктике

привлекателен для жителей европейской части Арктики, а Москва – для жителей азиатской. Исключениями стали Усинск и Новая Земля (предпочтение отдается Москве) в европейской части российской Арктики, а также Норильск (Санкт-Петербургу) – в азиатской [10].

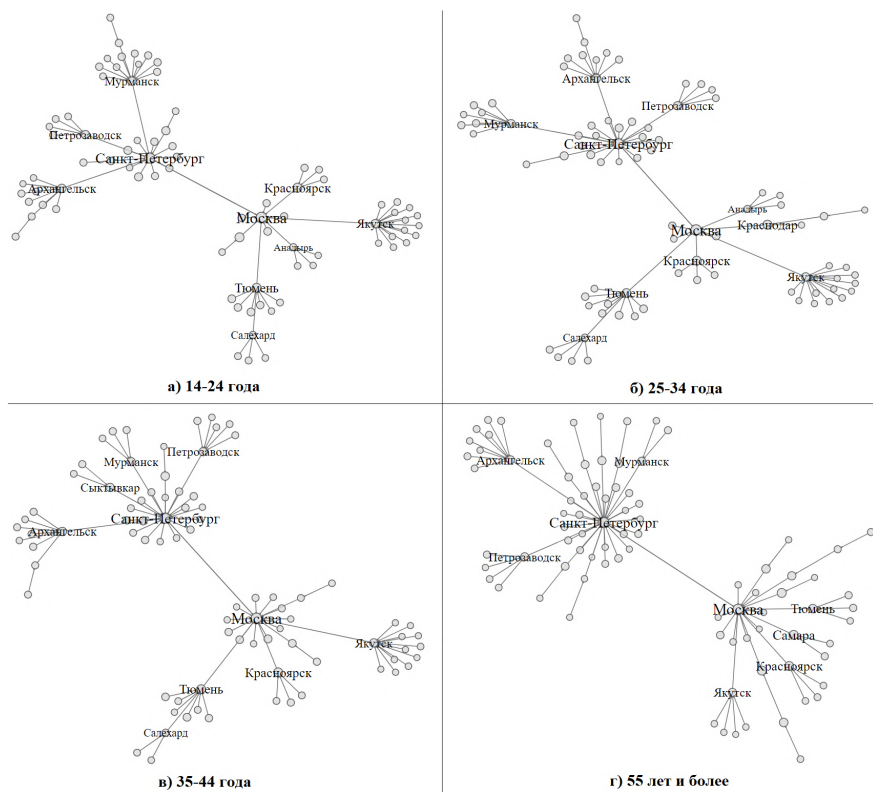


Рис. 2. Основные маршруты межмуниципальной миграции в российской Арктике по возрастным группам

Для 52 муниципальных образований, преимущественно в сельской местности, наиболее распространенным направлением перемещений стал их региональный административный центр. В Ямало-Ненецком АО районы разделились между двумя центрами – Салехардом и Тюменью, с преобладанием Тюмени. Из девяти региональных центров внутри российской Арктики расположены только пять (Архангельск, Мурманск, Салехард, Нарьян-Мар и Анадырь). Таким образом, большие города и региональные центры выполняют функцию миграционных хабов,

связывающих федеральные центры с арктическими территориями. Самый длинный миграционный маршрут до Москвы на полученном графе составляет 4 перемещения: Заполярный район => Нарьян-Мар => Архангельск => Санкт-Петербург => Москва. При этом средняя длина пути составила 2,17 перемещения.

Хотя маршруты миграций в различных возрастных группах похожи, между ними есть некоторые отличия (рис. 2). Москва наиболее привлекательна для людей среднего возраста, уже завершивших обучение. Крупные студенческие центры – Тюмень, Якутск, Мурманск – не привлекательны для представителей старших возрастов. В то же время, в старших возрастах намного выше удельный вес населения, перемещающегося в Санкт-Петербург.

Выявленные закономерности подчеркивают, что между арктическими городами существуют различия не только в интенсивности миграций по возрастным группам, но и в предпочтительных направлениях перемещений.

Заключение

По мере цифровизации различных сфер экономики и общественной жизни возрастает объем накапливаемых в цифровой среде данных о населении. Они становятся важным источником информации о демографических процессах и тенденциях. Как показало исследование, новые источники данных в ряде случаев позволяют получить информацию, более детальную, чем дает официальная статистика, и более масштабную, чем ту, что можно получить путем социологических исследований. При этом следует принимать во внимание преимущества и недостатки новых и традиционных источников данных, по возможности использовать их совместно. В дальнейших исследованиях планируется разработать методику построения демографических прогнозов с использованием новых цифровых источников данных.

Список источников

1. Смирнов А. В. *Цифровое общество: теоретическая модель и российская действительность* / А. В. Смирнов. DOI: 10.14515/monitoring.2021.1.1790 // *Мониторинг общественного мнения: экономические и социальные перемены*. 2021. № 1. С. 129-153.

2. Golder S. A. *Digital footprints: opportunities and challenges for online social research* / S. A. Golder, M. W. Macy. DOI: 10.1146/annurev-soc-071913-043145 // *Annual Review of Sociology*. 2014. V. 40 (1). P. 129-152.

3. Lazer D. *Data ex Machina: introduction to big data* / D. Lazer, J. Radford. DOI: 10.1146/annurev-soc-060116-053457 // *Annual Review of Sociology*. 2017. V. 43 (1). P. 19-39.

4. Kitchin R. *Big Data, New Epistemologies and Paradigm Shifts* / R. Kitchin. DOI: 10.1177/2053951714528481 // *Big Data & Society*. 2014. V. 1 (1). P. 1-12.

5. Богданов М. Б. *Возможности и ограничения цифровых следов и методов машинного обучения в социологии* / М. Б. Богданов, И. Б. Смирнов. DOI: 10.14515/monitoring.2021.1.1760 // *Мониторинг общественного мнения: экономические и социальные перемены*. 2021. № 1. С. 304-328.

6. Hughes C. *Inferring Migrations: Traditional Methods and New Approaches based on Mobile Phone, Social Media, and other Big Data* / C. Hughes, E. Zagheni, G. Abel [et al.]. DOI: 10.2767/61617. Brussels: European Commission. 2016. 41 p.

7. Замятина Н. Ю. *Виртуальная география виртуального населения* / Н. Ю. Замятина, А. Д. Яшунский. DOI: 10.14515/monitoring. 2018.1.07 // *Мониторинг общественного мнения: экономические и социальные перемены*. 2018. № 1. С. 117-137.

8. Смирнов А. В. *Население мировой Арктики: динамика численности и центры расселения* / А. В. Смирнов. DOI: 10.37482/issn2221-2698.2020.40.270 // *Арктика и Север*. 2020. № 40. С. 270-290.

9. Лыткина Т. С. *Вытеснение на Российском Севере: миграционные процессы и неолиберальная политика* / А. В. Смирнов, Т. С. Лыткина. DOI: 10.17238/issn2221-2698.2019.37.94 // *Арктика и Север*. 2019. № 37. С. 94-117.

10. Фаузер В. В. *Миграции населения российской Арктики: модели, маршруты, результаты* / В. В. Фаузер, А. В. Смирнов. DOI: 10.25283/2223-4594-2020-4-4-18 // *Арктика: экология и экономика*. 2020. № 4 (40). С. 4-18.

References

1. Smirnov A. V. *Cifrovoye obshchestvo: teoreticheskaya model' Irossijskaya dejstvitel'nost'* [Digital society: theoretical model and russian reality]. DOI: 10.14515/monitoring.2021.1.1790. *Monitoring obshchestvennogo mneniya: ekonomicheskiei social'nye peremeny*, 2021, no. 1, pp. 129-153. (In Russ.).

2. Golder S. A., Macy M. W. *Digital footprints: opportunities and challenges for online social research*. DOI: 10.1146/annurev-soc-071913-043145. *Annual Review of Sociology*, 2014, vol. 40 (1), pp. 129-152.

3. Lazer D., Radford J. *Data ex Machina: introduction to big data*. DOI: 10.1146/annurev-soc-060116-053457. *Annual Review of Sociology*, 2017, vol. 43 (1), pp. 19-39.

4. Kitchin R. *Big Data, new epistemologies and paradigm shifts*. DOI: 10.1177/2053951714528481. *Big Data & Society*, 2014, vol. 1 (1), pp. 1-12.

5. Bogdanov M. B., Smirnov I. B. *Vozmozhnosti i ogranicheniya cifrovyyh sledov i metodov mashinnogo obucheniya v sociologii [Opportunities and limitations of digital traces and machine learning methods in sociology]*. DOI: 10.14515/monitoring.2021.1.1760. *Monitoring obshchestvennogo mneniya: ekonomicheskie i social'nye peremeny*, 2021, no. 1, pp. 304-328. (In Russ.).

6. Hughes C. Zagheni E., Abel G. [et al.]. *Inferring Migrations: Traditional Methods and New Approaches based on Mobile Phone, Social Media, and other Big Data*. DOI: 10.2767/61617. Brussels: European Commission, 2016, 41 p.

7. Zamyatina N. Yu., Yashunsky A. D. *Virtual'naya geografiya virtual'nogo naseleniya [Virtual geography of virtual population]*. DOI: 10.14515/monitoring.2018.1.07. *Monitoring obshchestvennogo mneniya: ekonomicheskie i social'nye peremeny*, 2018, no. 1, pp. 117-137. (In Russ.).

8. Smirnov A. V. *Naselenie mirovoj Arktiki: dinamika chislennosti i centry rasseleniya [The arctic population: dynamics and centers of the settlement system]*. DOI: 10.37482/issn2221-2698.2020.40.270. *Arktikai Sever*, 2020, no. 40, pp. 270-290. (In Russ.).

9. Lytkina T. S., Smirnov A. V. *Vytesnenie na Rossijskom Severe: migracionnye process i neoliberal'naya politika [Expulsions in the Russian North: migration processes and neoliberal policy]*. DOI: 10.17238/issn2221-2698.2019.37.94. *Arktika i Sever*, 2019, vol. 37, pp. 94-117. (In Russ.).

10. Fauzer V. V., Smirnov A. V. *Migracii naseleniya rossijskoj Arktiki: modeli, marshruty, rezul'taty [Migration of the Russian Arctic population: models, routes, results]*. *Arktika: ekologiya i ekonomika*, 2020, vol. 4 (40), pp. 4-18. DOI: 10.25283/2223-4594-2020-4-4-18. (In Russ.).

Научное издание

IX Информационная школа молодого ученого

Сборник научных трудов

Рекомендовано к изданию
Ученым советом
Федерального государственного бюджетного учреждения науки
Центральной научной библиотеки Уральского отделения
Российской академии наук

Ответственный за выпуск *Н. А. Парменова*

Корректор: *Е. Л. Михайлова*
Компьютерная верстка *Е. Н. Иванова*

ISBN 978-5-8295-0788-6



Подписано в печать 29.11.2021. Формат 60x84 1/16.
Печать офсетная. Усл. печ. л. 32,3.
Тираж 300 экз. Заказ № 7240.

Книга издана согласно предоставленным материалам

ООО «Издательство УМЦ УПИ»
620000, Екатеринбург, ул. Гагарина, 35а, оф. 2

Отпечатано в типографии
ООО «Издательство УМЦ УПИ»
620000, Екатеринбург, ул. Гагарина, 35а, оф. 2