ENGLISH
SELECTION
2022

THE ELECTRONIC JOURNAL

DEMOGRAPHIC REVIEW

HSE UNIVERSITY

ISSN 2409-2274

FOUNDED BY ANATOLY G. VISHNEVSKY (1935-2021) IN 2014.

Released quarterly. Published since 2014.

All manuscripts are obligatory peer-reviewed.

Editorial office position does not necessarily coincide with the views of the authors.

The journal is registered on October 13, 2016 in the Federal Service for Supervision of Communications, Information Technology, and Mass Media.
Certificate of Mass Media Registration ЭЛ № ФС77-67362.
ISSN 2409-2274

Editorial address

Bolshoy Trekhsvyatitelskiy lane 3, office 303, Moscow, 109028, Russia
Phone: 8-495-772-95-90 * 11864 / *11824
www.demreview.hse.ru
demreview@hse.ru

Original papers

Spatial patterns of age-specific sex ratios in Russian intraregional migration

Aleksandr Gerasimov

THE ORIGINAL ARTICLE IN RUSSIAN WAS PUBLISHEDIN DEMOGRAPHIC REVIEW IN 2022, 9(1), 91-107.
https://doi.org/10.17323/demreview.v9i1.14575

Demographic features of industrial serfs in the 18th-19th centuries: parish registers of the village of Kudinovo in the Bogorodsky district, 1777-1862

Vlas Ryazanov

THE ORIGINAL ARTICLE IN RUSSIAN WAS PUBLISHEDIN DEMOGRAPHIC REVIEW IN 2022, 9(2), 4-21.
https://doi.org/10.17323/demreview.v9i2.16203

Digital traces of the population as a data source on migration flows in the Russian Arctic

Andrey Smirnov

THE ORIGINAL ARTICLE IN RUSSIAN WAS PUBLISHEDIN DEMOGRAPHIC REVIEW IN 2022, 9(2), 42-64.
https://doi.org/10.17323/demreview.v9i2.16205

Excess summer mortality at young ages: do long school holidays matter?

Polina Kuznetsova, Tatyana Maleva

THE ORIGINAL ARTICLE IN RUSSIAN WAS PUBLISHEDIN DEMOGRAPHIC REVIEW IN 2022, 9(2), 81-101.
https://doi.org/10.17323/demreview.v9i2.16207

Who needs (or doesn't) three years of parental leave?

Irina Kalabikhina, Polina Kuznetsova

THE ORIGINAL ARTICLE IN RUSSIAN WAS PUBLISHEDIN DEMOGRAPHIC REVIEW IN 2022, 9(3), 24-43.
https://doi.org/10.17323/demreview.v9i3.16468

Assessment of reliability of results of the 2019 Belarus population census based on an analysis of changes in the ethnolinguistic composition of the population

Aleksandr Sokolov

THE ORIGINAL ARTICLE IN RUSSIAN WAS PUBLISHEDIN DEMOGRAPHIC REVIEW IN 2022, 9(4), 61-103.
https://doi.org/10.17323/demreview.v9i4.16744

Демографическое обозрение / Demographic Review
2022;9(2) 42-64 DOI: https://doi.org/10.17323/demreview.v9i2.16205

Digital traces of the population as a data source on migration flows in the Russian Arctic

Andrey V. Smirnov
(av.smirnov.ru@gmail.com), Institute for Socio-Economic and Energy Problems of the North, Komi Science Centre of the Ural Branch of the Russian Academy of Sciences, Russia.

Abstract: *The digitalization of the economy and public life has expanded the possibilities of studying the population using digital traces – information that accumulates in the digital environment. Using digital traces, the article explores the migration of the population of the Russian Arctic, a huge macro-region that has experienced a significant outflow of population over the past decades. The text summarizes the experience of using digital traces in demographic research and formulates their strengths and limitations. Data from several digital platforms were used to study the population of the Russian Arctic. An analysis of the profiles of users of the social network VK.com made it possible to study the migration movements of the population of the Russian Arctic, and the data of the ticket service Tutu.ru provided information on air and rail movements. Using network analysis methods, the author studied migration and transport flows in the Russian Arctic at the municipal level. The article defines the features of migration and transport networks in the Arctic: low density, large distances between nodes, high relative mobility with small volumes of movements in absolute terms, a high proportion of hubs in migration exchange. The author identifies migration hubs and clusters, and migration flows are classified according to the directions of movement and types of municipalities. The text shows that the connectivity of the Arctic territories among themselves remains low, and the positive migration balance is mainly in regional capitals or cities outside the Arctic. The results of the study will improve the understanding of migration processes in the North and the Arctic, as well as the quality of demographic forecasts through more accurate modeling of migration flows.*

Keywords: *digitalization, digital traces, social networks, migration, transport network, migration flows, the Russian Arctic.*

Funding: *The article has been prepared with the support of the Russian Science Foundation, the project No 21-78-00081 "Development of tools for studying demographic processes in the context of the digitalization of society (in the case of the Russian Arctic)".*

For citation: *Smirnov A. (2022). Digital traces of the population as a data source on migration flows In the Russian Arctic. Demographic Review, 9(2), 42-64. https://doi.org/10.17323/demreview.v9i2.16205*

Introduction

In the modern world, almost all types of human activity are reflected in the Internet space. Therefore, digital data sources that preserve the history of human interaction with the Internet environment are increasingly becoming the subject of social science reflection (Katzenbach, Bächle 2019). Scientists analyze digital traces – "footprints" of human activity in the digital space, such as search queries, profiles and messages on social media, purchase information and data from global positioning systems (Dudina 2021: 5). New data sources provide information about society on a "massive and microscopic" scale at the same time (Golder, Macy 2014: 131). When studying migration, they make it possible to examine an enormous number of migratory flows with a high degree of detail and to reveal underlying patterns (Smirnov 2022).

In this article, the object of study is the spatial mobility of the population of the Arctic Zone of the Russian Federation, a macro-region in the north of the country which has already lost a third of its population over the past 3 decades, primarily as a result of interregional migration outflow (Fauzer, Smirnov 2020). Population decline persists today, but its study is hampered by the extreme unevenness and mosaic nature of demographic phenomena in the Arctic. The purpose of the study is to identify the spatial patterns of migration of the population of the Russian Arctic using digital traces. Along with migration, it will look at transport passenger flows, which characterize the connectivity of the Arctic territories with each other and with cities outside the Arctic zone. The information base of the study, in addition to official statistics, is made up of data from digital platforms: search engines, social media and ticketing services. To process these data, we used the methods of network analysis, a tool for studying the relationship between objects of any nature (Danchev, Porter 2021) - in the case of migration, between the territories of departure and arrival of people.

The article starts by systematizing the experience of using digital traces in demographic research and summarizing their advantages and disadvantages. Then, using new digital data sources, it looks at migration in the Russian Arctic at the regional and municipal levels. By analyzing migration and transport networks it reveals patterns in the movements of the population. At the end of the article, conclusions are drawn about the redistribution of the human potential of the Russian Arctic and some unresolved scientific problems are formulated.

Digital traces as a source of demographic data

Digital traces are the results of social interaction via digital tools and spaces as well as digital records of other culturally relevant materials (Cesare et al. 2018: 1980). The revolution in the use of digital traces by science has occurred thanks to the transition from small to large data. As a result of this transition, "the generation of data is continuous, exhaustive to a system, fine-grained, relational, and flexible across a range of domains" (Kitchin 2021: 61). Big data is being accumulated both in government information systems and on private digital platforms. The state collects data related to registration at the place of residence, employment, payment of taxes, visits to medical institutions, and receipt of various public services. Transport and utilities have become a digital network, equipped with numerous digital sensors that record people's behavior. Data from various sources is accumulated within large digital ecosystems covering many areas of life. A striking example is the Unified Portal of Public Services of the Russian Federation, which already has over 100 million users (Smirnov 2021: 148). As a result of the implementation of the national project "The Digital Economy" in Russia, we can expect an even greater deepening of the digitalization of society and public administration. In particular, by 2024 it is planned to transfer 70% of information systems and resources of federal authorities to a single cloud platform[1].

In addition to government information systems, big data is generated by commercial companies such as mobile phone operators (location, app use), travel and accommodation sites (reviews), social media (opinions, photos, personal info, location), transport providers (routes, traffic flows), website owners (clickstreams), financial institutions and retail chains (purchases), private surveillance and security firms (location, behavior) (Kitchin 2021: 62). Accumulation of data allows technology companies to extract greater profits through the vertical and horizontal integration of digital platforms. “In the twenty-first century advanced capitalism came to be centred upon extracting and using a particular kind of raw material: data. <...> Like oil, data are a material to be extracted, refined, and used in a variety of ways. The more data one has, the more uses one can make of them.” (Srnicek 2020: 37).

Big data is also generated through crowdsourcing and research projects. In Russia, the “Research Data Infrastructure” (RDI, data-in.ru) project should be noted, within the framework of which many valuable datasets have been prepared, for example, on the number of voters by precinct election commissions with reference to geographic coordinates in 2020. We will demonstrate this data set using the example of the Russian Arctic (Figure 1). The maps show that the data accumulated from the Internet pages of the election commissions, on the whole, quite accurately reflect the settlement system. Moreover, they evaluate the distribution of the population within cities in more detail - by polling stations.

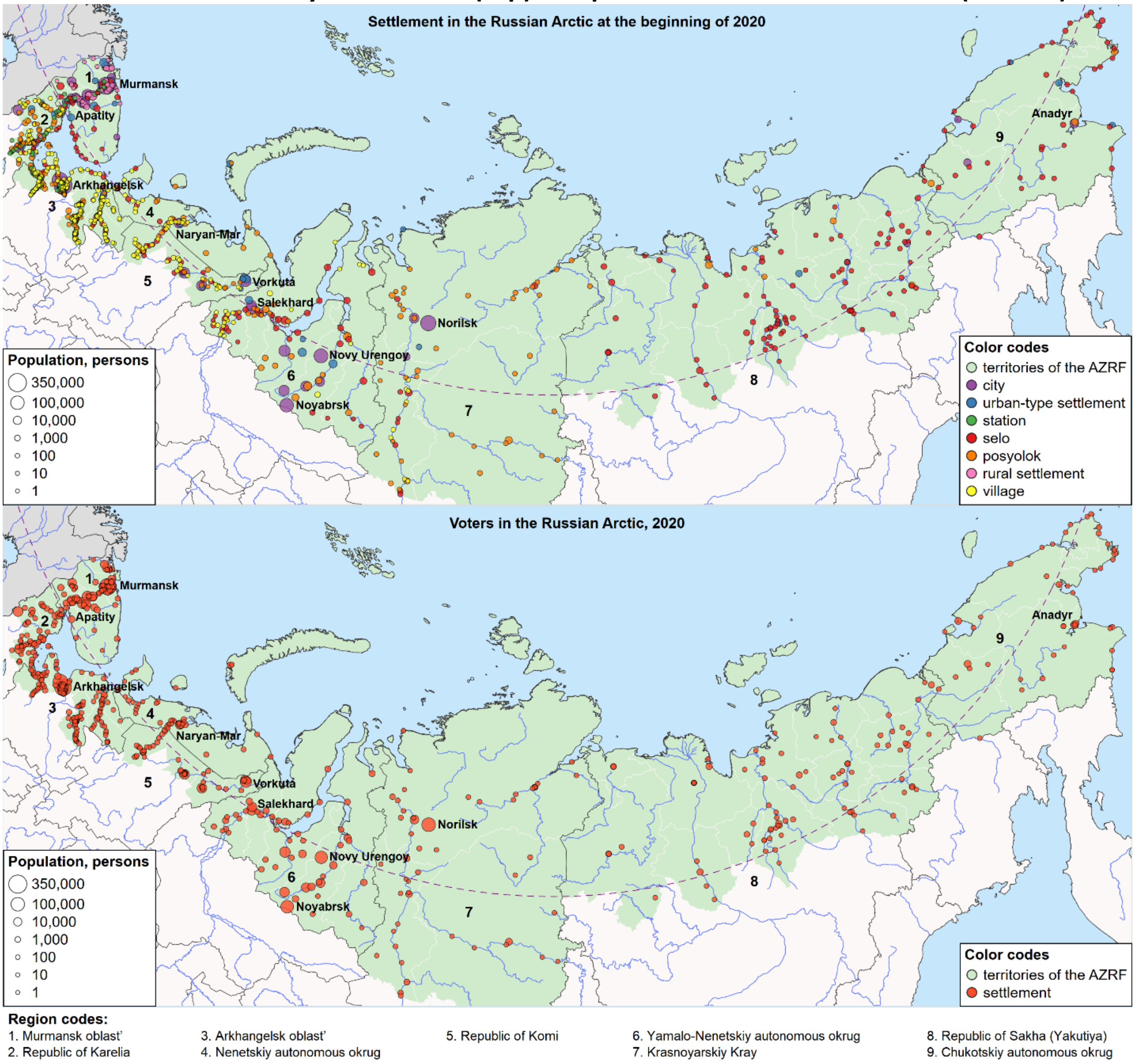
The advantages of digital data sources include large geographical coverage, continuous generation, and speed of collection and processing. They provide data on controversial topics because they are less biased by respondents' choice of socially acceptable responses than traditional surveys (Cesare et al. 2018: 1981). D. Leiser and J. Radford identify 3 types of digital data according to their sources and how they are acquired: digital life (the capturing of digitally mediated social behaviors), digital traces (only a record of the action, not the action itself) and digitized life (the movement of intrinsically analog behavior into digital form). Thanks to the development of analysis methods, research is increasingly using not only numerical and textual data obtained from the digital space, but also images, audio and video (Lazer, Radford 2017: 21-22, 33).

The disadvantages of new digital data sources include low representativity, fragmentation, vulnerability to changes, the possibility of algorithm errors, the presence of false information, bots and spam accounts, low reliability, duplication of information and limited access to data (Golder, Macy 2014; Lazer, Radford 2017). Digital data is difficult to interpret when not generated for research purposes, and digital concepts and terms may differ from theoretical ones. The problem of underrepresentation can be partly addressed by post-stratification or other bias correction techniques (Hughes et al. 2016). Often research questions are formulated in such a way that data correction is not required at all, for example, when the object of study is a virtual rather than a real population. Methods are being developed to reconcile digital trace data with other data sources, including census microdata (Alburez-Gutierrez et al. 2019). Digital research also faces many ethical challenges (Taylor, Floridi, van der Sloot 2017). To ensure data confidentiality, protect users from possible discrimination, and achieve reproducibility of studies,

[1] Passport of the federal project "Information infrastructure". https://files.data-economy.ru/Docs/Pass_Infrastructure.pdf

special open algorithms for collecting and processing digital data are being developed (Cesare et al. 2018: 1985).

Figure 1. Settlement in the Russian Arctic according to Rosstat, medical information and analytical centers (top) and precinct election commissions (bottom)



Source: Compiled by the author based on RDI datasets (http://data-in.ru/data-catalog/datasets/160; http://data-in.ru/data-catalog/datasets/203) using Natural Earth geodata (https://www.naturalearthdata.com/).

The development of methodological possibilities of research in connection with the introduction of digital data leads to attempts to develop theoretical optics suitable for new tools for the social sciences. It is proposed to turn digital traces into an independent object of study, and to rethink the problems of connection between micro- and macrolevels on the basis of D. Boullier's replication theory, which goes back to the works of G. Tarde (Dudina 2021). Applying the monadology of Leibniz to social phenomena, G. Tarde argues that "all phenomena are nebulous clouds resolvable into the actions emanating from a multitude of agents" (Tarde 2016: 32). According to Tarde, these figures do not have a coordinating center. D. Boullier distinguishes 3 stages in the development of sociological methods: the first source of data was statistics and censuses, then public opinion polls, and now digital traces. According to Boullier, digital traces reflect replications (repetitions, copying) of actions, ideas and practices (Boullier 2017). At the same time, digital platforms are perceived as a kind of "replicating machines", "allowing the dissemination of digital traces and making them available for research" (Dudina 2021: 5). Thus, the independent research value of digital traces is postulated. G. Igntatow comes to similar valuable conclusions from a practical point of view, reflecting on the theoretical foundations of the analysis of digital texts. He proposes to consider discourses as real emerging social phenomena, which allows them to be analyzed by rigorous methods (Igntatow 2016: 108).

The impact of new digital data sources on the social sciences is seen as revolutionary by many scholars (Kitchin 2014; Ledford 2020). In demography, digital traces began being used relatively recently, but they are already being used to solve a wide range of problems. Thus, images of cars from street panoramas are used to assess the socio-demographic characteristics of areas (Gebru et al. 2017). Valuable data is extracted from search engines and social media (McCormick et al. 2017; Zagheni, Weber, Gummadi 2017). Texts published by users on the web are used to analyze reproductive, self-preservation, matrimonial and migratory behavior. They can be studied both by frequency methods, using keywords, and by machine learning methods that can classify texts and highlight their semantic content and emotional coloring. For example, using automatic extraction and analysis of the opinions of social media users, it is possible to study various aspects of the reproductive behavior of the population (Kalabikhina et al. 2021).

Digital platforms are particularly useful when national statistics are unreliable for some processes (Cesare et al. 2018) or populations are studied to which access is difficult or costly (Edelmann et al. 2020). Mobile phone data and message geotagging are used to track the spatial mobility of the population (Hughes et al. 2016). For example, they were used to study compliance with self-isolation measures in various regions during the COVID-19 pandemic (Petrov et al. 2021: 9). The movements of specialists and scientists can be analyzed by the CVs and affiliations of scientific publications (Sudakova 2020). It has been shown that search query data can be used for short-term forecasts of fertility trends (Billari, D'Amuri, Marcucci 2013) and outbreaks of morbidity and mortality during a pandemic (Ahmad, Flanagan, Staller 2020). Digital traces are often used in conjunction with official statistics and the results of sociological research, complementing each other.

Characteristics of the Russian Arctic. Methods and data

The Arctic zone of the Russian Federation for 2022 includes 75 city okrugs and municipal districts[2] in 9 northern regions of Russia (Figure 1). At the beginning of 2022, 2592.9 thousand people lived in the Russian Arctic (excluding the results of the 2021 population census). The Arctic zone accounts for about 30% of the area, 1.8% of the population and 6% of Russia's gross regional product. The key sector of the economy is resource extraction. About 90% of Russia's natural gas, a significant share of oil, coking coal, and non-ferrous metals are produced in the Arctic. Economic specialization, spatial remoteness, and uncomfortable climatic conditions have an impact on demographic structures (Heleniak, Bogoyavlenskiy 2014; Fauzer, Lytkina 2017; Zamyatina, Yashunsky 2017). Since the population of the Arctic is relatively young and lives mainly in cities, it is characterized by high values of digitalization indicators. More than 90% of

[2] Federal Act of 13 July 2020 entitled "State support for entrepreneurship in the Arctic zone of the Russian Federation".

the population from 15 to 74 years old are active Internet users[3], which corresponds to the level of the most developed countries in the world. Therefore, digital traces can be a fairly representative source of data on Arctic populations.

Official statistics on the size and migration of the population of the Arctic territories were obtained from the Unified Interdepartmental Statistical Information System[4] (UISIS) and the Database of Municipal Indicators[5] (DMI) of Rosstat. Digital traces of the population were analyzed using three data sources: the Yandex service "Keyword Statistics", the project "Virtual Population of Russia" and the Tutu.ru service data set. Let's look at their capabilities and limitations.

Yandex Keyword Statistics[6] is a service for assessing user interest in topics. It allows you to get information about the popularity of a particular query in the search engine in the context of regions. The names of the regions of the Arctic were entered as queries, which made it possible to assess the interest that Yandex users in some regions had in other regions. The presence of search queries does not guarantee that users are planning to move to or visit a region. Nevertheless, interest in a region can characterize the intensity of cultural, social or economic interactions.

The project "Virtual Population of Russia"[7] was implemented with the support of the Russian Geographical Society and the Keldysh Institute of Applied Mathematics of the RAS. It contains geo-referenced data for January-March 2015 from user profiles of the most popular social network in Russia at that time, VK. The data of the project make it possible to analyze migratory movements at the regional and municipal levels by age groups, as well as friendships between people. Of the 88 million accounts that indicate a place of residence or a place of study, 9 million have more than one place of residence, which makes it possible to analyze migration flows. The migration dataset takes into account only the chronologically last change of location for each user. The set's limitations include the availability of data for only one point in time, as well as the fact that people tend to indicate not the municipality where they actually live, but the nearest large city to it (Zamiatina, Yashunsky 2018). In addition, users are less likely to record short-term and return migrations.

The dataset of the Tutu.ru ticketing service[8] about traveling around the country was created to predict the spread of the coronavirus infection COVID-19[9]. It contains information on the number of movements between cities by plane, train and bus[10] in April 2019. The number of passengers is not limited to the number of tickets sold through the Tutu.ru service, but has been restored to 100%. According to the developers of the dataset, buses are the most inaccurate part of the dataset due to the presence of “grey” carriers. For the Arctic settlements, there are only 11 routes (mainly from Arkhangelsk and Petrozavodsk). Therefore, in this study, we restrict

---

[3] Statistical information on socio-economic development of the Arctic zone of the Russian Federation. Rosstat. https://rosstat.gov.ru/storage/mediabank/arc_zona.html
[4] Unified interdepartmental statistical information system. Rosstat and Ministry of Digital Development of Russia. https://www.fedstat.ru/
[5] Database of municipal indicators. Rosstat. https://www.gks.ru/dbscripts/munst/
[6] Keyword Statistics. Yandex. https://wordstat.yandex.ru/
[7] Virtual population of Russia. http://webcensus.ru/
[8] Dataset Tutu.ru and data from Open Data Science model. https://story.tutu.ru/dataset-tutu-ru-i-dannye-modeli-open-data-science/
[9] Scenarios of infection in specific cities based on the dataset of movement of people in Russia. Habr. https://habr.com/ru/company/tuturu/blog/494700/
[10] Covid19-tutu. GitHub. https://github.com/ods-ai-ml4sg/covid19-tutu/blob/master/data/raw_data.csv

ourselves to data on aircraft (558 routes) and trains (712 routes). Air and rail data also have limitations. In the eastern part of the country, market coverage is worse. This applies in particular to helicopter routes between the cities of the Far East and the propeller-driven aviation of Yakutia.

Migration and transport flows were studied using network analysis methods (Danchev, Porter 2021). Network science tools are used to study migration both at the interregional (Maier and Vyborny 2008) and intercountry (Danchev and Porter 2018) levels. A network is a set of nodes and a set of links (edges) between them. Migration and transport networks are most conveniently represented as directed and weighted. The directions of the links correspond to the directions of movement, from the place of departure to the place of arrival, while the weights correspond to the number of people who have moved. Six networks were built. Three of them include only links with at least one node in the Arctic: migration, air and railway passenger flows. Three more similar networks were built according to data for the whole of Russia and used for comparative analysis.

The limitations of the study include the fact that the data sources used are not in all cases synchronous. They refer to the period from 2015 to 2022. It should also be noted that migration and traffic flows characterize different types of migration: official data from Rosstat that of long-term migration, and data from ticketing services that of any movements, including short-term ones for recreation, medical treatment, education, work or family. Estimating rotational migration, which is important for the Arctic, is complicated by the fact that rotational workers are difficult to separate from other passengers, and travel to the work site may be not on regular, but on charter flights, on which data are less available.

Using the NetworkX package in the Python programming language, algorithms for calculating the main characteristics of migration and transport networks were developed. To search for communities (clusters) in networks, an asynchronous label propagation algorithm (Raghavan, Albert, Kumara 2007) was used, based on the idea that related nodes usually belong to the same cluster. Network visualization was carried out using the GraphPlot.jl and Graphs.jl packages in the Julia language. Network visualization algorithms tend to place nodes in such a way that connected nodes are close to each other and the number of link intersections is minimized. For networks with more than one cluster, a modified force-directed graph drawing algorithm of the Fruchterman-Reingold was used (Fruchterman, Reingold 1991). For the network of aviation passenger flows, in which it is difficult to distinguish clusters, the stress majorization graph drawing algorithm was used (Gansner, Koren, North 2004). The VegaLite.jl package was used to create chart maps, and the chorddiag package in the R programming language was used to build a chord diagram.

Connectivity of the Arctic territories: digital data and official statistics

Connectivity is understood as the degree of connection (link) of cities or territories, expressed through the presence and quantity of migration, transport movements, or other interactions between them. First, we will look at the connectivity of the Arctic territories with each other at the regional level (Figure 2). We will analyze four indicators, three of them based on digital traces of the population and the last one on official statistics:

- the number of friendships per person of the virtual population of the region with residents of another region according to the social network VK in January-March 2015. The indicator is not symmetrical. Although the two regions have the same number of friendships in both directions, their virtual populations differ. Therefore, for example, one virtual resident of Karelia has more friendships with residents of the Murmansk region than vice-versa;

- the number of migration movements per 1,000 people of the virtual population, according to the data of the VK social network. The chronologically last change of residence according to user accounts for January-March 2015 is used;
- the popularity among a region's residents of a search query with the name of another region in the Yandex search engine. Regional popularity is a region's share of displays for a given query divided by the share of all search results displayed for that region. If the regional popularity exceeds 1, there is a high interest in this request in the region; if it is less than 1, it is low;
- the number of migration movements per 1,000 people of the real population by regions of departure and arrival according to UISIS data. The average annual value for 2015-2021 is used.

Figure 2. Some indicators of the connectivity of the Arctic regions

	a) Number of friendships per person of the virtual population of the region (Virtual Population of Russia, 2015)									
	connected region									
analyzed region	№	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Murmansk obl.	(1)		0.55	0.51	0.01	0.12	0.02	0.14	0.04	0.00
Rep. Karelia	(2)	0.77		0.31	0.00	0.10	0.02	0.13	0.01	0.00
Arkhangelsk obl.	(3)	0.42	0.18		0.16	0.36	0.03	0.15	0.08	0.00
Nenetskiy AO	(4)	0.21	0.09	6.50		1.38	0.17	0.13	0.02	0.00
Rep. Komi	(5)	0.13	0.08	0.51	0.05		0.04	0.13	0.01	0.00
Yamalo-Neneetskiy AO	(6)	0.05	0.03	0.09	0.01	0.10		0.16	0.02	0.00
Krasnoyarskiy Kray	(7)	0.05	0.04	0.07	0.00	0.04	0.02		0.07	0.00
Rep. Sakha (Yakutia)	(8)	0.06	0.01	0.15	0.00	0.02	0.01	0.30		0.00
Chukotskiy AO	(9)	0.07	0.03	0.08	0.00	0.03	0.02	0.16	0.13	

b) Number of migration movements per 1000 persons of the virtual population (Virtual Population of Russia, 2015)									
region of arrival									
region of departure (1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	№
	1.54	0.84	0.01	0.17	0.04	0.30	0.10	0.01	(1)
2.09		0.34	0.01	0.09	0.03	0.11	0.02	0.00	(2)
1.30	0.48		0.30	1.15	0.08	0.19	0.04	0.00	(3)
0.97	0.44	24.6		3.30	1.31	0.24	0.05	0.00	(4)
0.31	0.16	0.87	0.10		0.19	0.19	0.03	0.00	(5)
0.09	0.03	0.10	0.02	0.27		0.29	0.01	0.00	(6)
0.16	0.05	0.08	0.00	0.05	0.06		0.11	0.01	(7)
0.22	0.07	0.45	0.01	0.09	0.11	1.79		0.02	(8)
0.77	0.51	0.84	0.00	0.51	0.64	3.02	1.22		(9)

	c) Regional popularity of search query naming other region (Yandex Keyword statistics, 13.04.2022)									
	requested region									
analyzed region	№	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Murmansk obl.	(1)		2.64	1.11	0.99	0.68	0.72	0.13	0.50	0.68
Rep. Karelia	(2)	3.35		0.78	0.75	0.63	0.55	0.08	0.57	0.88
Arkhangelsk obl.	(3)	1.20	1.09		4.79	1.43	0.87	0.10	0.54	0.84
Nenetskiy AO	(4)	1.08	0.71	11.1		8.92	4.95	0.13	1.19	1.85
Rep. Komi	(5)	0.52	0.57	1.35	2.25		1.38	0.10	0.68	0.75
Yamalo-Neneetskiy AO	(6)	0.27	0.43	0.37	1.08	0.87		0.23	0.94	2.13
Krasnoyarskiy Kray	(7)	0.18	0.29	0.19	1.27	0.43	0.74		1.03	0.88
Rep. Sakha (Yakutia)	(8)	0.27	0.35	0.21	0.80	0.89	0.68	0.36		2.46
Chukotskiy AO	(9)	0.62	0.69	0.45	3.26	0.49	3.38	0.57	6.41	

d) Number of migration movements per 1000 persons of real population of region of departure (Rosstat, average for 2015-2021)									
region of arrival									
region of departure (1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	№
	1.50	1.22	0.02	0.17	0.04	0.17	0.03	0.02	(1)
1.72		0.27	0.01	0.07	0.02	0.05	0.00	0.01	(2)
1.04	0.19		0.35	0.55	0.05	0.05	0.01	0.01	(3)
0.40	0.06	9.49		2.13	0.16	0.16	0.04	0.00	(4)
0.24	0.08	0.82	0.12		0.17	0.07	0.02	0.01	(5)
0.07	0.03	0.10	0.01	0.22		0.25	0.04	0.01	(6)
0.05	0.01	0.02	0.00	0.02	0.05		0.16	0.02	(7)
0.02	0.01	0.01	0.00	0.01	0.03	0.56		0.02	(8)
0.31	0.18	0.11	0.00	0.16	0.17	0.85	0.40		(9)

Source: Compiled by the author based on data from the Virtual Population of Russia project, Yandex (dated April 13, 2022) and Rosstat.

In general, regions that are geographically close to each other are more connected, as evidenced by the green colors of many cells located near the diagonal of the tables. All connectivity indicators reach the highest value between the Nenets Autonomous Okrug and the Arkhangelsk Oblast, of which it is a part. Low connectivity values are recorded in Yakutia, where the population is predominantly rural, with no large cities (an exception is the connection between Yakutia and the Krasnoyarsk Territory). The linear correlation between migration according to official statistics and social network data is 0.851 (n = 71, links from the Nenets Autonomous Okrug to the Arkhangelsk region are excluded as statistical outliers). The correlation between migration according to statistics and friendships has a high value – 0.789. The correlation of migration with the intensity of search queries is significantly lower – 0.542.

It can be concluded that, although in absolute terms migration rates obtained from official statistical databases and digital traces of the population differ, they show similar patterns. The connectivity of territories in the digital environment usually means that in reality there is indeed a high connectivity (migration, social, cultural) between them. Moving on to the next level of detail, let's consider the networks of intermunicipal movements and their main characteristics (Table 1).

Table 1. Networks of migration and passenger movements of the population of the Russian Arctic and Russia

Indicator	Migration		Air transport		Rail transport	
	AZRF	Russia	AZRF	Russia	AZRF	Russia
Size of network (number of nodes)	2112	2201	85	173	160	574
Number of links	32199	334529	558	2951	712	12125
Strongly connected	no	no	no	no	no	no
Weakly connected	yes	yes	yes	yes	yes	no
Reciprocity	0.421	0.419	0.828	0.855	0.775	0.760
Average link weight	6.1	11.0	978.1	2071.6	384.2	689.9
Average distance of movement, km	…	…	1707.7	1632.9	956.3	601.3
Average degree	30.5	304.0	13.1	34.1	8.9	42.2
Average strength	186.9	3319.1	12842.1	70674.0	3419.8	29146.4
Network density	0.007	0.069	0.078	0.099	0.028	0.037
Average path length	1.875	1.936	1.925	2.246	2.489	2.236 *
Diameter of network **	4	3	5	5	6	5 *
Heterogeneity parameter	14.675	2.475	3.157	2.581	3.934	3.409
Degree_assortativity_coefficient	-0.519	-0.202	-0.635	-0.364	-0.644	-0.213
Clustering coefficient	0.525	0.468	0.302	0.507	0.398	0.640
Number of clusters when decomposing by label propagation method	2	1	1	1	4	6

Source: Compiled by the author based on data from webcensus.ru and Tutu.ru.

*Note: * - to calculate the value, the network was converted to weakly connected by removing 5 nodes related to an isolated section of the railway on Sakhalin; ** - to calculate the diameter, all networks were converted to non-directional. AZRF - Arctic zone of the Russian Federation.*

In terms of the size and number of links, migration networks lead, since they are not limited by the capabilities of the transport infrastructure and can connect any settlements. Of the transport networks, rail is larger than air, since populated areas have fewer airports than railway stations. However, in the Arctic, due to its remoteness, the proportion of cities with airports is higher than in the country as a whole. The Russian railway network is not connected, as it includes an isolated railway on Sakhalin Island. When calculating some indicators, this railway was not included. In transport networks, the indicator of reciprocity is higher, i.e., movements are more often carried out in both directions.

The average link weight (average number of moves per flow) is higher in air networks due to there being fewer possible routes. In air travel, the average travel distance is also higher than by rail - for the Arctic, almost twice as high (1707.8 versus 956.3 km). The average degree of a node shows the average number of links or neighbors of nodes throughout the network. In the Arctic migration network, the average degree is an order of magnitude lower than in the Russian migration network. In transport networks, the Arctic's lag is also large due to its remoteness from the main centers of settlement. This is also evidenced by the lower values of density of the Arctic networks. Since transport and migration networks are weighted, it is possible to calculate a weighted degree for them - the node strength. The patterns are similar and even more pronounced.

The shortest path is the minimum number of links that must be made along the path connecting 2 nodes. The average path length is calculated by averaging the lengths of the shortest paths over all pairs of nodes. The lowest values of the average path length are observed in migration networks, due to the large number of links. In Arctic networks, the diameter (the length of the longest shortest path) is greater than or equal to the value for the whole of Russia. The largest diameter (6 movements) is in the railway network, and the smallest (4) in the migration network.

The more hubs (nodes with a higher degree) in the network, the higher the heterogeneity parameter. In the Arctic, heterogeneity is higher and is highest in migration networks. Hubs are Moscow, St. Petersburg and the administrative centers of the regions. However, in other networks too the heterogeneity parameter is quite high. There are hubs in all six networks; they will be discussed in more detail below. In all networks, the assortativity coefficient is negative. This suggests that high-degree nodes (hubs) are more often connected to low-degree ones. Moreover, this dependence is manifested more clearly in the Arctic networks.

The clustering coefficient shows the proportion of pairs of neighbors of a node that are connected to each other. In the Arctic, it is the highest in migration networks. As a rule, either hubs or settlements located close to each other are more connected. In air transport, the network clustering coefficient is relatively low. The maximum number of clusters is in railway networks (4 in the Arctic and 6 in Russia). The composition of the clusters will be discussed below.

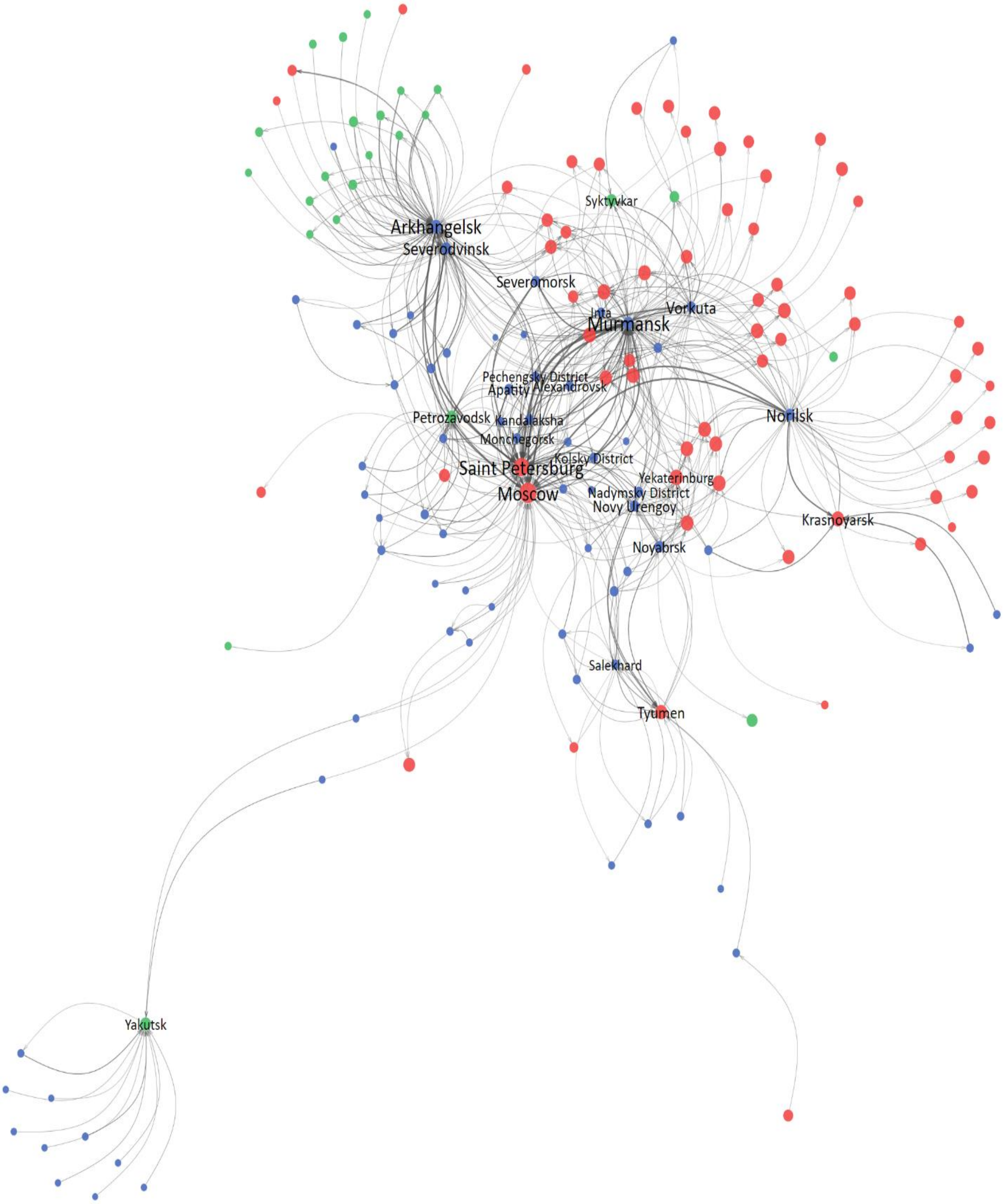
Network analysis showed differences in the patterns of migration and transport movements between Russia and its Arctic part in almost all indicators. The main reason for the differences is the low density of connections, due to the spatial remoteness of the macroregion, as well as the high influence of hubs - the most significant network nodes. Let's focus attention on them.

Migration flows and hubs in the Russian Arctic

Clusters (communities) are sets of nodes with a higher density of connections within than between sets. Two clusters in the network of migration movements have been highlighted (Figure 3). The first includes some rural areas of Yakutia (located in the lower left part of the figure), and the second, the remaining municipalities.

The largest hub in the Arctic migration network is Murmansk (Table 2). It accounts for 17.3% of movements. It is followed by Arkhangelsk (12.9%), St. Petersburg (12.7%), Moscow (8.9%) and Norilsk (8.0%). Moreover, the migration balance is much better in the federal capitals: in St. Petersburg and Moscow, incoming flows greatly exceed outgoing ones. In most Arctic cities, according to digital traces, there persists a negative migration balance. Vorkuta (-7.5 thousand movements) and Norilsk (-7.3 thousand) especially stand out.

Figure 3. Network of intermunicipal migrations in the Russian Arctic



Source: Compiled by the author based on data from the Virtual Population of Russia project.

Note: Only flows larger than 50 people are reflected. Municipal entities of the Arctic zone are marked in blue, the Far North of Russia (except for the Arctic) is marked in green, and other regions of Russia are marked in red. The thickness and brightness of the line is proportional to the size of the flow, the size of the circle is proportional to the population of the municipality.

Table 1. Characteristics of the largest nodes of the network of migration movements of the Russian Arctic

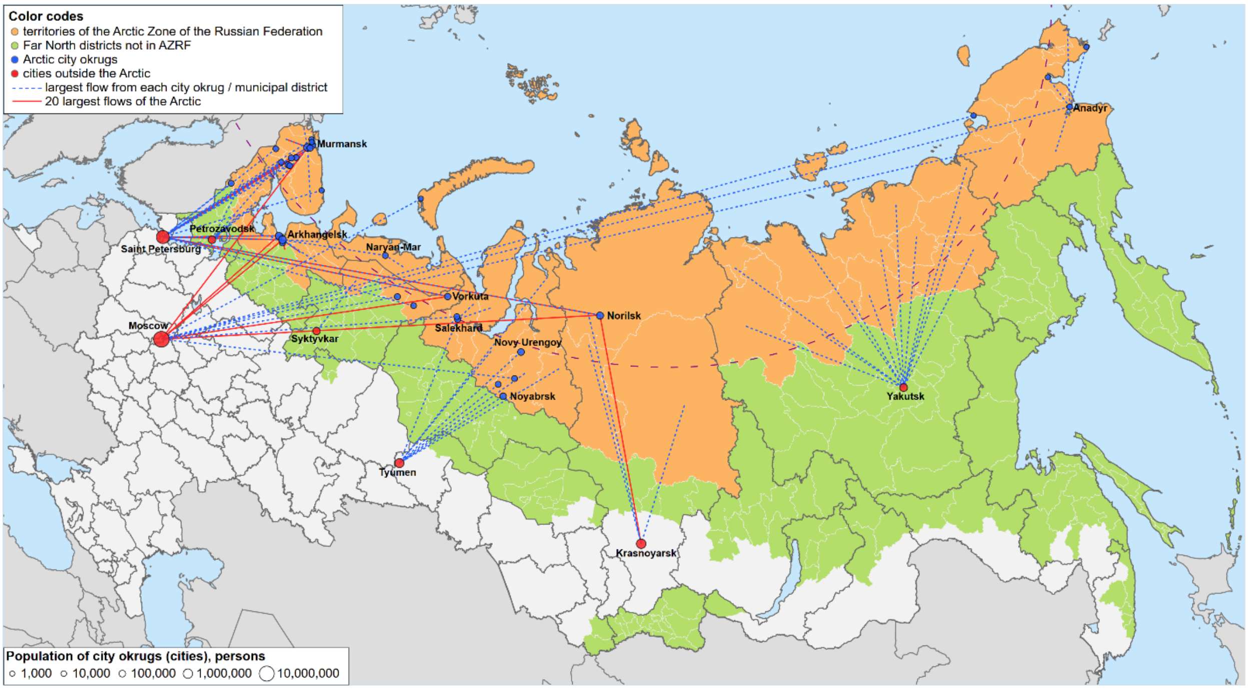
№	City okrug / municipal district	Outflows		Inflows		Balance	
		quantity, units	size, persons	quantity, units	size, persons	quantity, units	size, persons
1	Murmansk	810	16.235	1.421	18.109	611	1.874
2	Arkhangelsk	611	11.167	853	14.473	242	3.306
3	St. Petersburg	64	2.945	74	22.287	10	19.342
4	Moscow	68	3.308	75	14.483	7	11.175
5	Norilsk	753	11.655	985	4.349	232	-7.306
6	Severodvinsk	488	6.292	668	5.122	180	-1.170
7	Vokruta	749	9.283	603	1.734	-146	-7.549
8	Novy Urengoy	396	3.971	911	3.706	515	-265
9	Severomorsk	374	4.671	723	2.636	349	-2.035
10	Noyabrsk	340	3.288	801	2.901	461	-387
11	Apatity	378	3.199	397	1.880	19	-1.319
12	Nadymskiy district	362	3.408	601	1.644	239	-1.764
13	Aleksandrovsk	362	3.234	617	1.728	255	-1.506
14	Krasnoyarsk	35	745	60	3.492	25	2.747
15	Monchegorsk	308	2.650	418	1.264	110	-1.386
16	Inta	416	3.249	237	555	-179	-2.694
17	Petrozavodsk	36	1.130	48	2.607	12	1.477
18	Tyumen	26	485	49	3.239	23	2.754
19	Pechengskiy district	345	2.693	404	937	59	-1.756
20	Usinsk	334	2.113	477	1.515	143	-598
21	Salekhard	231	1.553	490	1.993	259	440
	Total	32.199	199.052	32.199	199.052	0	0

Source: Compiled by the author based on data from the Virtual Population of Russia project.

Note: Ranked in descending order of total movements.

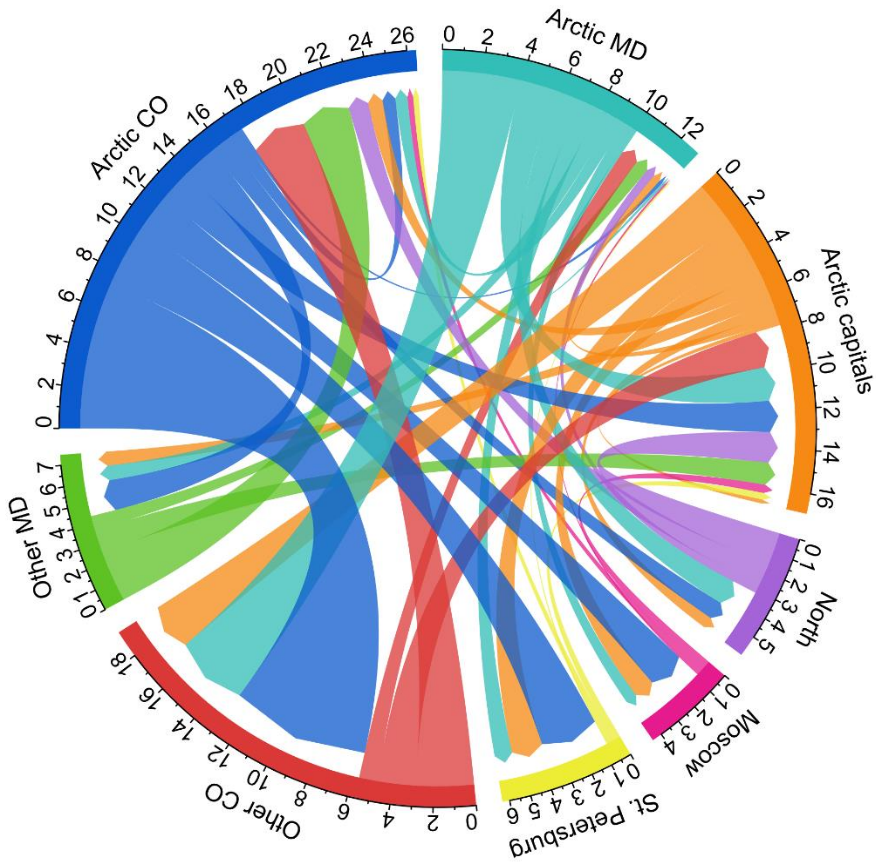
Of the 20 largest flows, only 5 link municipalities both of which are located in the Arctic: from the Kola district and Severomorsk to Murmansk, from Severodvinsk to Arkhangelsk and back, from the Pinezhsky district to Arkhangelsk. Another 1 connects the Arctic with the municipality of the Far North - from the Kholmogorovsky district to Arkhangelsk (Figure 4). The rest connect the Arctic city okrugs with Moscow, St. Petersburg and Krasnoyarsk. Residents of the European part of the Arctic more often move to St. Petersburg, while those in the Asian part move to Moscow (exceptions are Norilsk, Usinsk and Novaya Zemlya). Residents of city okrugs are more likely to move to Moscow and St. Petersburg, while residents of municipal districts are more inclined to move to regional capitals (Figure 5). This may be due to the availability of relocation resources for residents of cities and urban-type settlements, most often specializing in the resource extraction industry or the transportation of natural resources. The administrative centers of the regions located in the Arctic (Arkhangelsk, Murmansk, Salekhard, Naryan-Mar and Anadyr) give about the same number of people as they get. But while people arrive mainly from the northern and Arctic municipalities, they leave for cities outside the Arctic.

Figure 4. The largest migration flows in the Russian Arctic



Source: Compiled by the author based on data from the Virtual Population of Russia project using Natural Earth geodata (https://www.naturalearthdata.com/).

Figure 5. Migration flows of the Russian Arctic by groups of municipalities, %



Source: Compiled by the author based on data from the Virtual Population of Russia project.

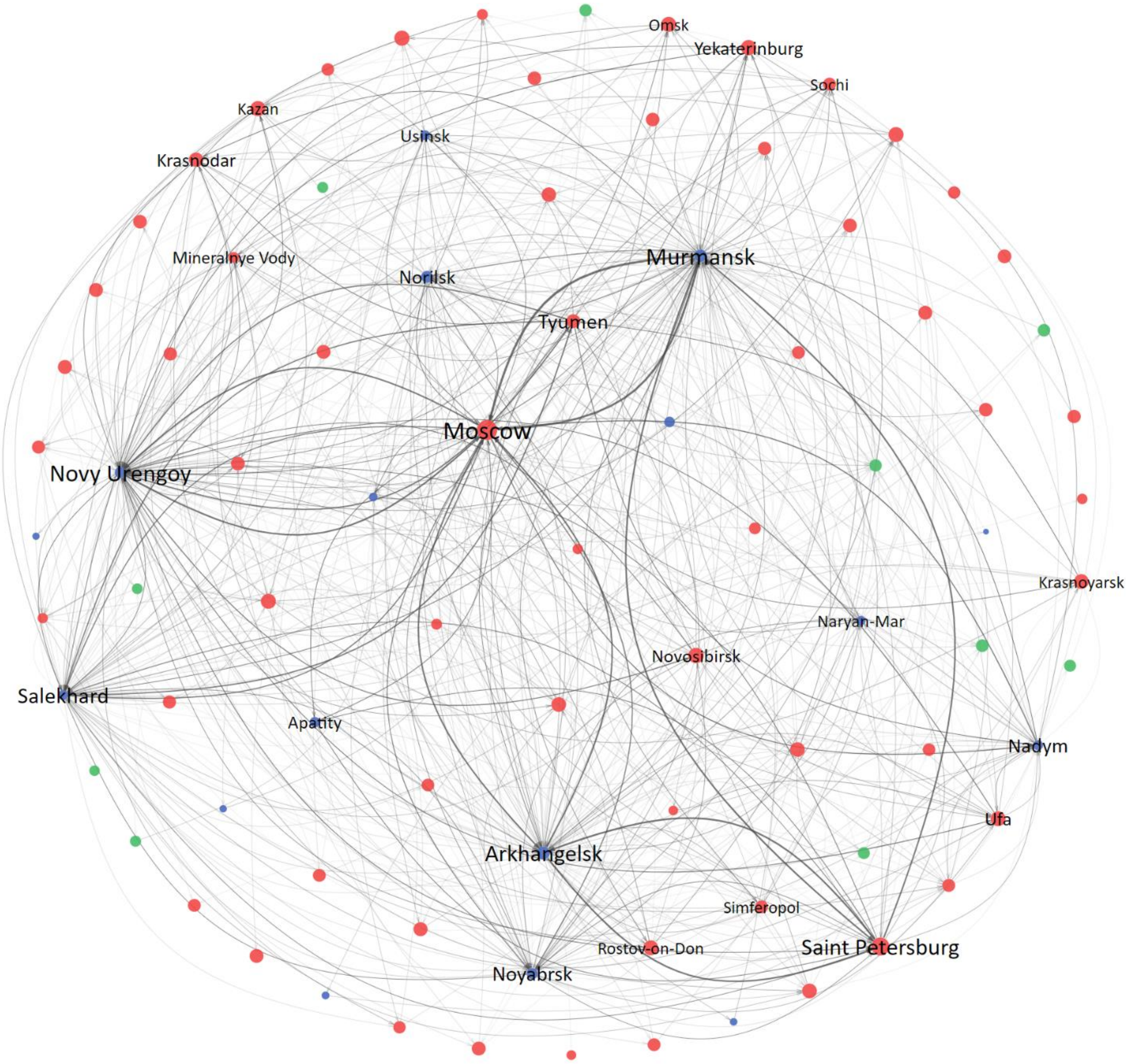
Note: CO – city okrug, MD - municipal district.

Numerical data confirm the main migration trends in the Arctic recorded by official statistics (Fauzer, Smirnov 2020) and reveal them with a high degree of detail. The population is becoming concentrated in a small number of large cities and their environs. Population decline from older resource cities and rural areas remains high.

Transport flows and mobility of the Arctic population

It was not possible to identify explicit clusters in the network of air travel movements. Almost all parts of the Arctic have similar patterns of movement with a high proportion of outflows to Moscow and back (Figure 6).

Figure 6. Air passenger traffic network in the Russian Arctic

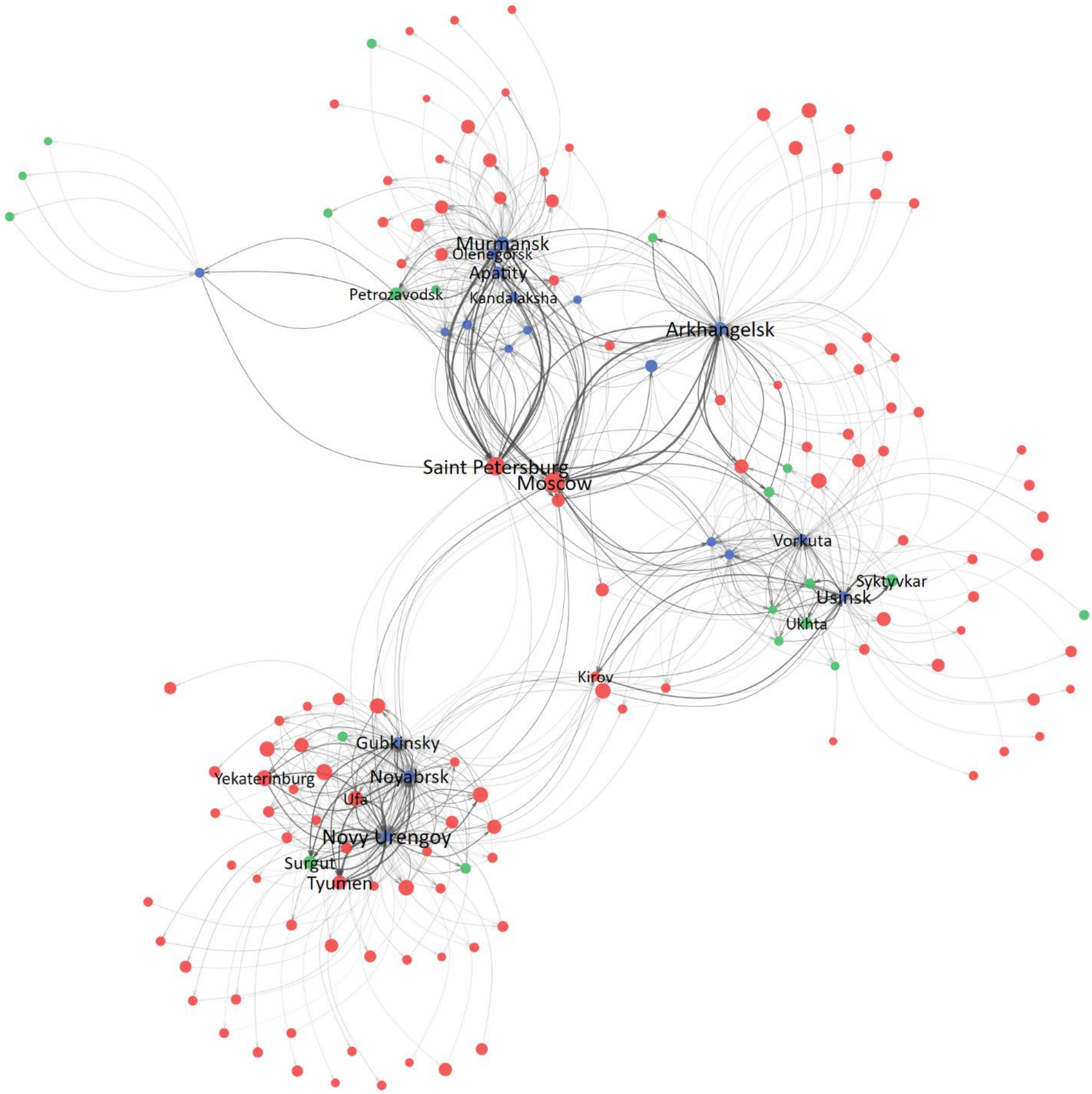


Source: Compiled by the author according to the Tutu.ru service.

Note: Settlements in the Arctic Zone are marked in blue, in the Far North of Russia (except for the Arctic) in green, and in other regions of Russia in red.

In the network of railway movements, clusters are distinguished that roughly correspond to the branches of the Russian Railways (RZD): Oktyabrskaya, Severnaya, Sverdlovskaya (Figure 7). Moreover, the Severnaya Railway is divided into 2 clusters corresponding to branches to Arkhangelsk and Vorkuta. An intermediate position between the cores of the clusters is occupied by Moscow, St. Petersburg and Kirov, whose railways are not included in the listed branches of Russian Railways and are used as transfer hubs.

Figure 7. Rail passenger traffic network in the Russian Arctic



Source: Compiled by the author according to the Tutu.ru service.

Note: Settlements in the Arctic Zone are marked in blue, in the Far North of Russia (except for the Arctic) in green, and in other regions of Russia in red.

In the air travel network, 15.1% of movements are made from Moscow, and another 17.7% - to Moscow (Table 3). Thus, the capital accounts for a third of all movements. In addition to Moscow, there are several other major hubs: Murmansk (25.4% of movements), Novy Urengoy

(22.8%), St. Petersburg (17.4%), Arkhangelsk (16.9%) and Salekhard (12.8%). In the railway network, the share of large hubs is lower. Novy Urengoy accounts for 19.6% of movements, Moscow and Arkhangelsk - 16.5% each, St. Petersburg - 16.0%. The shares of Murmansk (12.9%), Usinsk (11.6%) and Noyabrsk (11.1%) are high. The large passenger turnover of Novy Urengoy is associated, among other things, with rotational migrations.

Table 3. Characteristics of the largest nodes of the network of air and rail passenger traffic in the Russian Arctic

№	City	Outflows		Inflows		Balance	
		quantity, units	size, persons	quantity, units	size, persons	quantity, units	size, persons
				Air transport			
1	Moscow	13	82,372	13	96,451	0	14,079
2	Murmansk	52	76,607	47	62,057	-5	-14,550
3	Novy Urengoy	45	68,159	48	56,435	3	-11,724
4	St. Petersburg	12	47,251	12	47,531	0	280
5	Arkhangelsk	40	43,975	40	48,169	0	4,194
6	Salekhard	42	37,810	33	32,313	-9	-5,497
7	Tyumen	7	21,830	8	24,302	1	2,472
8	Noyabrsk	26	18,952	27	18,001	1	-951
9	Nadym	22	18,444	16	13,590	-6	-4,854
10	Norilsk	34	13,306	24	9,954	-10	-3,352
	Total	558	545,791	558	545,791	0	0
				Rail transport			
1	Novy Urengoy	56	29,643	49	23,982	-7	-5,661
2	Moscow	18	21,960	18	23,280	0	1,320
3	Arkhangelsk	38	22,898	35	22,156	-3	-742
4	St. Petersburg	15	22,047	15	21,630	0	-417
5	Murmansk	33	20,642	28	14,747	-5	-5,895
6	Usinsk	29	17,022	22	14,798	-7	-2,224
7	Noyabrsk	36	15,356	33	15,013	-3	-343
8	Tyumen	3	10,581	3	12,987	0	2,406
9	Gubkinskiy	33	11,983	31	8,835	-2	-3,148
10	Apatity	25	10,604	22	9,922	-3	-682
	Total	712	273,581	712	273,581	0	0

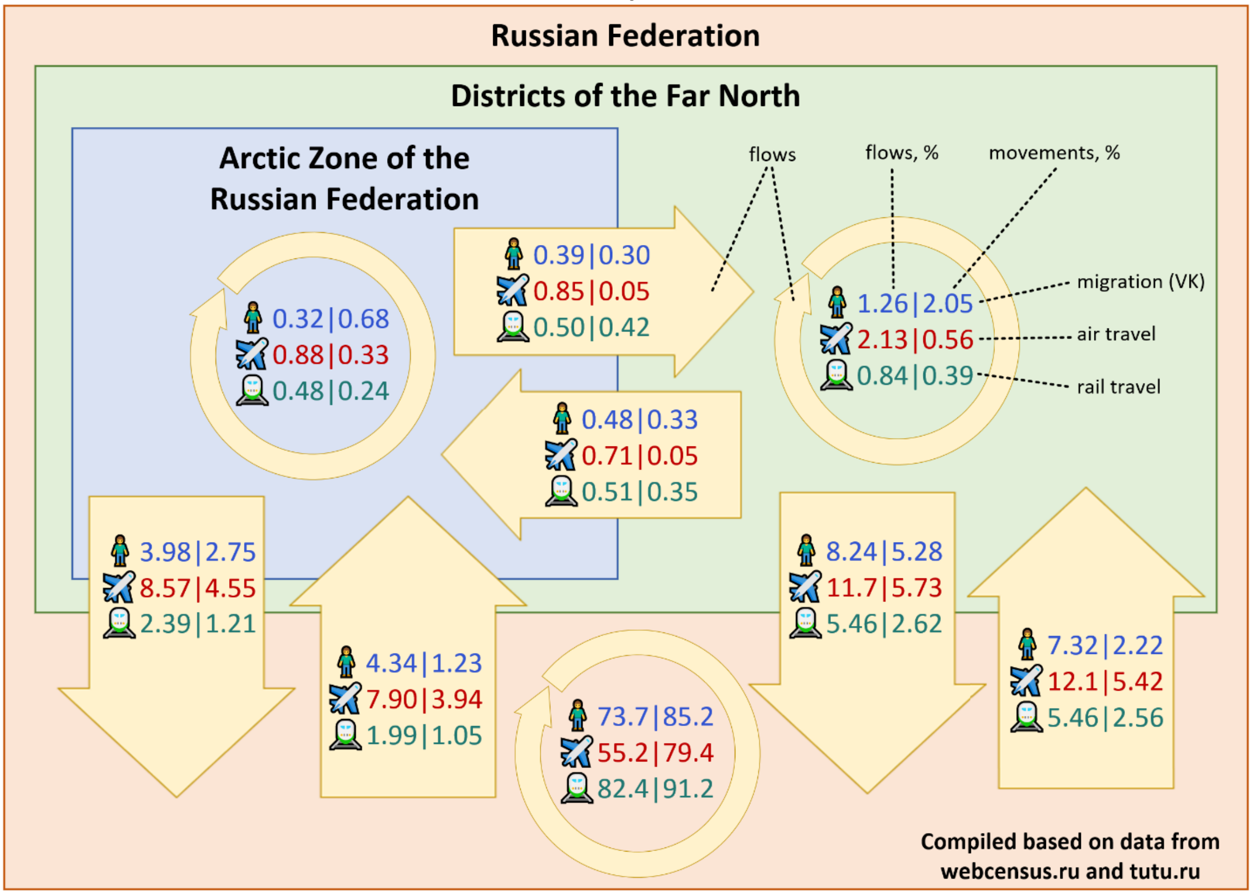
Source: Compiled by the author based on Tutu.ru data as of April 2019.

Note: Ranked in descending order of total moves.

The difference in the number of incoming and outgoing flows from the same nodes is explained by the short analysis time period (1 month), which could lead to underestimation of rare flights; a lack of data on small flows (less than 10 people for rail transport and less than 50 for air); and the presence of compound routes, including more than 1 movement with transfers. If the tickets were bought with different bookings or from different companies, it is very difficult to identify composite routes. This leads to an overestimation of the share of hubs in traffic flows.

The largest flows in air transport are between Murmansk and Moscow, and in rail transport, between Arkhangelsk and Moscow. None of the 20 largest air flows connect 2 Arctic cities. In rail transport, the situation is similar, but there are 5 flows that connect Arctic cities with cities located in the non-Arctic part of the Far North. These are flows from Usinsk to Syktyvkar and back, from Novy Urengoy to Surgut and back, from Noyabrsk to Surgut. Most of the largest flows connect the Arctic municipalities with the capital cities and administrative centers of the regions in the central and southern parts of the country. To consider the connectivity of the Arctic territories in more detail, let's combine all types of the examined migration and transport flows into one scheme (Figure 8).

Figure 8. Distribution of migration and transport flows by directions, % of the total number of flows/movements



Source: Compiled by the author based on data from the Virtual Population of Russia project and the Tutu.ru service.

Although only 1.8% of Russia's population lives in the Arctic, they account for 5.3% of migratory movements, 3.3% of rail movements, and 8.9% of air movements. Consequently, the Arctic population is more mobile, and air transport, which connects remote settlements with federal centers, is of particular importance. There are very few displacements within the Arctic (0.68% of migrations, 0.33% of air and 0.24% of railways), which indicates a low connectivity of the Arctic territories with each other. A much larger share is accounted for by movements between Arctic and non-Arctic settlements (4.6, 3.0 and 8.6% of all movements in the country, respectively). In this regard, the importance should be noted of completing the construction of such large infrastructure projects as the Northern Latitudinal Railway, Belkomur (White Sea – Komi Rep. – Ural railway) and Barentskomur (Barents Sea – Komi – Ural railway). At the same time, to ensure the connectivity of territories in the Arctic, it is necessary to expand the use of regional and local aviation.

Conclusion

The study showed that new data sources that have emerged due to the development of digital technologies make it possible to obtain detailed and timely data on migration processes, while network analysis provides suitable tools for systematizing and comprehending this information. In the territorial context, digital traces can reach the level of settlements or even contain the coordinates of individual places. They reflect different types of migration by duration, direction and reason. Data obtained from social media can contain very detailed socio-demographic characteristics of the population, and search queries and digital texts make it possible to analyze the migration intentions and preferences of residents without large time and material costs. Digital traces make it possible to include millions of Internet users in research, and at the same time to gather information at the micro level, including from remote and hard-to-reach areas. Processing methods, approaches to interpretation and ethical aspects of the use of digital data on the movements of people are still being developed, and further expansion of their explanatory and predictive potential can be expected.

The study of the digital traces of the population of the Russian Arctic made it possible to identify key migration and passenger flows in the macroregion. It has been confirmed that the connectivity of the territories of the Arctic with each other is quite low, and the bulk of the movements concern flows with cities outside the Arctic. It is shown that there are differences in the models of demographic behavior between residents of the Arctic capitals, the administrative centers of the regions, and of other urban districts and municipal districts, as well as between European and Asian territories. The migration and transport hubs of the Russian Arctic have been identified. Moscow and St. Petersburg account for more than a fifth of migration, a third of rail and half of air movements. Moreover, in the federal capitals, incoming migration flows are much larger than outgoing ones. The identification of clusters in migration networks showed a high degree of isolation of territories in the north of Yakutia, and in railway networks, the division of the network into four parts due to the limitations of the existing transport infrastructure. Together with the weak development of the motorway network in most of the Russian Arctic, this hinders the development of horizontal cooperation between residents and organizations in the Arctic Zone.

Since the demographic dynamics in the Arctic are determined primarily by migration flows, taking into account their characteristics at the district and settlement levels will make it possible to build more accurate forecasts of the size and composition of the population. In crisis situations, the speed of generating digital traces is of great importance. Thus, digital sources of data on population migration and morbidity make it possible, without waiting for the publication of official statistics, to develop forecasts for the development of the coronavirus pandemic and make management decisions based on them. The study presents a static picture of patterns of migratory movements, without taking into account the time factor. To conduct such a study in dynamics, the tools of temporal networks can be used. Particular attention in the future study of the problem should be given to seasonal fluctuations in migration, which are significant for remote areas with raw materials. In addition, digital traces can be used to study international and rotational migrations, which are of particular importance for the labor markets of the Arctic in the face of a declining resident population.

References

Dudina V.I. (2021). "Reassembling sociology": the digital turn and the search for new theoretical optics. *Sociological Studies*, 11, 3-11. https://doi.org/10.31857/S013216250016829-4.

Zamyatina N.Yu., Yashunsky A.D. (2018.). Virtual geography of the virtual population. *Monitoring of Public Opinion: Economic and Social Changes*, 1, 117-137. https://doi.org/10.14515/monitoring.2018.1.07.

Kalabikhina I.E., Lukashevpich N.V., Banin E.P., Alibaeva K.V., Reblei S.M. (2021). Automatic extraction of social network users' attitudes on reproductive behavior issues. *Program Systems: Theory and Applications*, 12:4(51), 33-63. https://doi.org/10.25209/2079-3316-2021-12-4-33-63.

Kitchin R. (2021). Data-driven, networked urbanism. In E. Lapina-Kratasyuk, O. Zaporozhets, A. Vozyanov (Eds.), *City Networks: People. Technologies. Authorities* (pp. 58-80). Moscow: New Literary Review.

Smirnov A.V. (2021). Digital society: theoretical model and Russian reality. *Public Opinion Monitoring: Economic and Social Changes,* 1(161), 129-153. https://doi.org/10.14515/monitoring.2021.1.1790.

Smirnov A.V. (2022). Digital Demography Methods for Forecasting Migration Processes. *Economy of the regions*, 18(1), 133-145. https://doi.org/10.17059/ekon.reg.2022-1-10.

Srnicek N. (2020). *Platform capitalism*. Moscow: Publishing House of the Higher School of Economics.

Sudakova A.E. (2020). Migration of scientists: the Digital footprint and scientometry. *Perspectives of Science and Education*, 3(45), 544-557. https://doi.org/10.32744/pse.2020.3.39.

Tarde G. (2016). *Monadology and sociology*. Perm: Giele Press.

Fauzer V.V., Lytkina T.S. (2017). Migration processes in the Russian North. *Social policy and sociology*. 16:1 (120), 141-149. https://doi.org/10.17922/2071-3665-2017-16-1-141-149.

Fauzer V.V., Smirnov A.V. (2020). Migration of the Russian Arctic population : models, routes, results. *The Arctic: Ecology and Economy*, 4(40), 4-18. https://doi.org/10.25283/2223-4594-2020-4-4-18.

Ahmad I., Flanagan R., Staller K. (2020). Increased internet search interest for GI symptoms may predict COVID-19 cases in US hotspots. *Clinical Gastroenterology and Hepatology*, 18(12), 2833-2834. https://doi.org/10.1016/j.cgh.2020.06.058.

Alburez-Gutierrez D., Aref S., Gil-Clavel S., Grow A., Negraia D.V., Zagheni E. (2019). Demography in the Digital Era: New data sources for population research. In: *SIS2019. Smart statistics for smart applications.* Milano: Pearson. https://doi.org/10.31235/osf.io/24jp7.

Billari F.C., D'Amuri F., Marcucci J. (2013). *Forecasting births using Google. Population Association of America Annual Meeting*. https://paa2013.princeton.edu/papers/131393.

Boullier D. (2017). Big data challenges for the social sciences: from society and opinion to replications. In *eSymposium*, 7(2), 1-17. https://www.boullier.bzh/wp-content/uploads/EBul-Boullier-Jul2017.pdf

Cesare N., Lee H., McCormick T., Spiro E., Zagheni E. (2018). Promises and pitfalls of using digital traces for demographic research. *Demography*, 55, 1979-99. https://doi.org/10.1007/s13524-018-0715-2.

Danchev V., Porter M.A. (2018). Neither global nor local: Heterogeneous connectivity in spatial network structures of World migration. *Social Networks*, 53, 4-19. https://doi.org/10.1016/j.socnet.2017.06.003

Danchev V., Porter M.A. (2021). Migration networks: applications of network analysis to macroscale migration patterns. In M. McAuliffe (Ed.), *Research handbook on international migration and digital technology* (pp. 70-90). Cheltenham: Edward Elgar Publishing. https://doi.org/10.4337/9781839100611.

Gansner E., Koren Y., North S. (2004). Graph Drawing by Stress Majorization. *Lecture Notes in Computer Science*, 3383, 239-250. https://doi.org/10.1007/978-3-540-31843-9_25.

Gebru R., Krause J., Wang Y., Chen D., Deng J. Aiden E.L., Fei-Fei L. (2017). Using deep learning and Google street view to estimate the demographic makeup of neighborhoods across the United States. *PNAS*, 114(50), 13108-13113. https://doi.org/10.1073/pnas.1700035114.

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

Edelmann A., Wolff T., Montagne D., Bail C. (2020). Computational Social Science and Sociology. *Annual Review of Sociology*, 46, 61-81. https://doi.org/10.1146/annurev-soc-121919-054621.

Fruchterman T.M.J., Reingold E.M. (1991). Graph drawing by force-directed placement. *Software: Practice and Experience*, 21(11), 1129-1164. https://doi.org/10.1002/spe.4380211102.

Heleniak T., Bogoyavlenskiy D. (2014). Arctic Populations and Migration. In Larsen J.N., Fondahl G., Rasmussen H. (Eds.), *Arctic Human Development Report. Regional Processes and Global Linkages* (pp. 53-104). Copenhagen: Nordic Council of Ministers. https://doi.org/10.6027/TN2014-567.

Hughes C., Zagheni E., Abel G., Wi´sniowski A., Sorichetta A., Weber I., Tatem A.J. (2016). *Inferring migrations: Traditional methods and new approaches based on mobile phone, social media, and other big data*. Luxembourg: Publications Office of the European Union. https://doi.org/10.2767/61617.

Igntatow G. (2016). Theoretical foundations for digital text analysis. *Journal for the Theory of Social Behaviour*, 46(1), 104-120. https://doi.org/10.1111/jtsb.12086.

Katzenbach C., Bächle T.C. (2019). Defining concepts of the digital society. *Internet Policy Review*, 8(4). https://doi.org/10.14763/2019.4.1430.

Kitchin R. (2014). Big Data, New Epistemologies and Paradigm Shifts. *Big Data & Society*, 1(1), 1-12. https://doi.org/10.1177/2053951714528481.

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

Ledford H. (2020). How Facebook, Twitter and other data troves are revolutionizing social science. *Nature*, 582, 328-330. https://doi.org/10.1038/d41586-020-01747-1.

Maier G., Vyborny M. (2008). Internal migration between US States: A social network analysis. In J. Poot, B. Waldorf, L.W. Wissen (Eds.), *Migration and Human Capital*. Cheltenham, UK: Edward Elgar Publishing. URL: https://www.econstor.eu/handle/10419/117573.

McCormick T.H., Lee H., Cesare N., Shojaie A., Spiro E.S. (2017). Using Twitter for demographic and social science research: tools for data collection and processing. *Sociological Methods & Research*, 46(3), 390-421. https://doi.org/10.1177/0049124115605339.

Petrov A.N., Welford M., Golosov N., DeGroote J., Degai T., Savelyev A. (2021). The “second wave” of the COVID-19 pandemic in the Arctic: regional and temporal dynamics. *International Journal of Circumpolar Health*, 80(1). https://doi.org/10.1080/22423982.2021.19254461.

Raghavan U.N., Albert R., Kumara S. (2007). Near linear time algorithm to detect community structures in large-scale networks. *Physical Review E*, 76(3). https://doi.org/10.1103/physreve.76.036106.

Taylor L., Floridi L., van der Sloot L. (Eds.). (2017). *Group privacy: New challenges of data technologies.* Cham, Switzerland: Springer. https://doi.org/10.1007/978-3-319-46608-8.

Zagheni E., Weber I., Gummadi K. (2017). Leveraging Facebook’s advertising platform to monitor stocks of migrants. *Population and Development Review*, 43(6178), 721-734. https://doi.org/10.1111/padr.12102.

Zamyatina N., Yashunsky A. (2017). Migration cycles, social capital and networks. A new way to look at Arctic mobility. In M. Laruelle (Ed.), *New Mobilities and Social Changes in Russia’s Arctic Regions* (pp. 59-84). London and New York, Routledge.