

УДК: 311.21:[316.774:004.738.5]
JEL Classification: C83
doi: 10.31767/nasoa.4.2019.01

Г. В. ГОЛУБОВА,
кандидат економічних наук, доцент,
доцент кафедри статистики,
Національна академія статистики, обліку та аудиту,
e-mail: g_kondrya@ukr.net,
ORCID: 0000-0003-4847-5235,
Researcher ID: I-6414-2018

Використання можливостей Інтернету для соціальних статистичних досліджень

В статті розглянуто особливості, проблеми і перспективи проведення соціальних статистичних досліджень з розвитком Інтернету і соціальних мереж. У найбільш розвинених країнах чисельність користувачів Інтернет майже дорівнює загальній чисельності населення, що створює безліч нових можливостей для статистичних соціальних досліджень. На відміну від традиційних методів досліджень, соціальні мережі містять величезні масиви даних про користувачів та їхню діяльність. Цю інформацію можна представити в цифровій формі, що уможливорює її подальше оброблення і проведення на її основі різних видів аналізу. При цьому головним завданням залишається розроблення методології, спеціально призначеної для соціальних досліджень в Інтернеті.

Ключові слова: методологія, соціальні дослідження, Інтернет, соціальні мережі, комунікації.

Г. В. ГОЛУБОВА,
кандидат экономических наук, доцент,
доцент кафедры статистики,
Национальная академия статистики, учета и аудита

Использование возможностей Интернета для социальных статистических исследований

В статье рассмотрены особенности, проблемы и перспективы проведения социальных статистических исследований с развитием Интернета и социальных сетей. В наиболее развитых странах численность пользователей Интернета почти равна общей численности населения, что создает множество новых возможностей для статистических социальных исследований. В отличие от традиционных методов исследований, социальные сети содержат огромные массивы данных о пользователях и их деятельности. Эту информацию можно представить в цифровой форме, что создает возможность для ее дальнейшей обработки и проведения на ее основе различных видов анализа. При этом главной задачей остается разработка методологии, специально предназначенной для социальных исследований в Интернете.

Ключевые слова: методология, социальные исследования, Интернет, социальные сети, коммуникации.

Using Internet Capacities for Social Statistical Surveys

The article describes the features, problems and benefits of social statistical surveys by means of Internet. The changes taking place with the development of the Internet and social networks create new challenges, in particular as regards the methodological support of such research.

In the most developed countries, the population using the Internet is almost equal to the total population, and this trend will continue to spread in other countries as well. This creates many new opportunities for statistical social research.

Unlike conventional observations, social networks exist in such a way that they contain a great deal of information about users and their activities, which can be digitized and presented as a database in which information or information will be constantly updated or accumulated. Thus, it makes it suitable for generalization, calculation, classification, measurement, as well as for a number of advanced statistical and other analyzes. It is theoretically and practically possible to study the population as a whole, and not just its sample population.

Data on social networks are available in a continuous time series and space, which allows for constant monitoring of trends and contributes to a deeper understanding of cause and effect changes. This approach improves the descriptive plane of research, unlike conventional observation methods. At the same time, there are some difficulties, including the reliability of the observation data. Because, it is quite difficult to check the reliability of the socio-demographic characteristics of the data obtained (gender, age, education, etc.). In view of this, in the international practice the method of correcting the results of research in social networks with national observation data is used.

Therefore, the key issue remains to be the development of special methodology for social statistical surveys in the Internet, which will take into account all the features and specificities of the Internet environment on the whole and its users in particular.

Keywords: *methodology, social research, Internet, social networks, communication.*

The development of information society is associated with the advancement of information and communication technologies. The powerful capacities of online social networks triggered change in the social information media through changing the manners and modes of social communications, reformatting business organization etc. New information and communication opportunities coming with the emergence of Internet call for statistical support for new approaches to performing social statistical studies, including their methodological backing.

Studies of social networks as a means of communication have been in focus of domestic and foreign researchers: T. Zbrytska [1], R. Zmysna [2], K. Dolzhenkova [3], A. Tabanova [1], C. Fedushko [2], M. Cattelan [4], D. Firth [4], A. Malley [9], J. Onnela [9], C. Varin [4] and others. They elaborate on the impact of social networks on the development of mass media, consumers and media services. Emphasis should be made on the specifics involved in performing social studies in the Internet environment, in order to see the relevance of the conventional methodology for statistical studies to the challenges of our time.

The article's objective is to investigate the peculiarities of the social networks' development in the context of social statistical studies.

The social networks are referred to as groups formed in Internet to associated the individuals who can have common interests or relations of friendship, or the individuals showing the need to publish or monitor the information published by others, to exchange information that may be useful for others, and to help each other by advice or any other way by answering questions put by other members of a network etc. A closely related notion is the notion of social or new mass media. Basically, it refers to an association (group of individuals), with the difference in the observance perspective: while in the social networks emphasis is made on relations and links between their members and the purpose of these

relations (from business to romantics), in new mass media it is made on publications and review of various content. So, the social networks are the social “infrastructure” of new media, whereas mass media represent dynamic and interactive exchange of ideas, knowledge, impressions, thoughts and opinions.

According to data from the State Statistics Services of Ukraine [12], the number of Internet users in Ukraine was 26.1 million persons in 2018, or 61.6 % of the total population, whereas in some countries it exceeds 90%: Denmark (97.3%), Luxembourg (97%), the U. K. (94.9%), the Netherlands (94.7%), Sweden (92.1%), Japan (91.3%), China (90.5%) etc. [8].

A most explicit peculiarity of social networks is communications. The latest years are marked by the increasing scopes of communications in social networks (mediated or virtual communications) by exchanging information though intermediation of various services in Internet, mobile or stationary phones etc. But it is obvious that the largest growth in the mediated communications is demonstrated by social networks, which is illustrated by Table 1.

Table 1

Number of minutes of personal communications and Internet communications

Year	Personal communications, min./day	Internet communications, min./day	Ratio, %
2008	221.5	11.2	19.8
2013	228.3	18.4	12.4
2018	213.9	38.6	5.5

Source: constructed by the author by data from [8].

Data from Table 1 show that the time spent for communications via Internet has been gradually and stably increasing. While at the end of 2008 “*tête-à-tête*” communications were nearly 20 times longer than the ones via Internet, in 2018 they were only 5.5 times longer. It should be noted that the growth in Internet communications is regarded by psychologists as an unfavorable social effect ousting “*tête-à-tête*” communications, but we treat social network participants as potential respondents, so the growing numbers of Internet users and overall scopes of the social network communications are regarded by us as a positive phenomenon.

Another peculiarity of Internet communications is their planetary dimension. It means that the communications are going on in online mode across the planet, while prior to these networks appearance they were possible only by use of telephone conversations, with far lesser scopes and, therefore, far narrower contents. An outstanding feature of the mediated communications is its multimedia aspect that involves inclusion of not only words or sounds, but sophisticated audiovisual matter, enabling one to have a clearer image of his/her interlocutor.

In fact, the emerged social networks gave birth to a new media environment integrating conventional mass media, i. e. printing, radio, TV, films, etc. So, the analogy with conventional mass media leads to the conclusion that theoretically each individual Internet user is a media, an object and an entity, representing an actual or potential producer, broadcaster, advertiser and audience as a receiver of transmitted information and consumer of information, goods, services, ideas, and even consumer of propagandist messages [11].

It can be asked if a social impact of social media really exists and if it is possible to study social phenomena in Internet environment. Unlike conventional communities, social networks exist in a way to give very significant “clues” on what are their users and what they do: their members write text messages, receive and display visual or music matter or even films. If we scrutinize the contents of these messages in social networks as indicators of potential social studies, we can make the conclusion that infinite numbers of facts available in Internet can give us a much better idea about the life of these population groups than conventional ways of research. It makes social networks applicable for potential social statistical studies, especially for descriptive purposes.

The increasing pace of information globalization requires revisions of the available statistical methodology with respect to social studies, because studies in Internet environment

have been rapidly expanding in parallel with ones conducted in conventional ways (direct contact through interviews).

The information about the real life, behaviors, tastes, hobbies or material well-off of people can now be digitalized and represented in form of a database by way of testing or simulating their behavior or disputes in focus groups. It enables a wider range of socially oriented topics than in conventional studies, which obviously go beyond the official statistical observations.

A social unit of study is Internet users, with the sample scope being unlimited and more informative because it contains extended information of physical and subject contents. Moreover, a sample can theoretically be approximated with the size of general population, but this does not require additional costs.

A distinct feature of such social studies in Internet is their continuity, unlike ad hoc surveys that are only capable to make a momentary snapshot. The monitoring of a studied object conducted in this way can result in a deeper understanding of causal relationships and effects.

For the methodology of social statistical studies it means that “benchmark element” of an empirical study is digital data about an observation unit, which can be easily identified, grouped, classified, measured and incorporated into an extended statistical or other analysis.

The author’s comparison of a conventional observation and a one conducted in Internet is summed up and shown in Table 2.

Table 2

A comparison of a conventional observation and a one conducted in social online networks

Aspect	Conventional observation	Internet observation
Observation way	Direct interview (“tête-à-tête”)	Indirect interview
Observation form	Specially organized observation (conventional interview, questioning etc.)	Specially organized observations with the extended factual account (text, audio, visual records etc.)
За ступенем охоплення даних	Incomplete, sample observation	Incomplete, sample observation; compete observation. Theoretically, all the general population can be studied
Time of facts record	Ad hoc or regular observation	Ad hoc, regular observation, continuing monitoring
Way of record	Record of data from the respondent words, questionnaires etc.	Self-record
Statistical tools	A set of statistical forms, instructions etc.	Electronic form
Description of data	The obtained data are not digitalized; their aggregation, grouping and summing up are required	The data obtained already have digital form; database. “Self-measurement” by use of tools designed by web browsers etc.
Reliability of data	Reliability of data can be texted	Reliability of data cannot be easily texted
Data control (arithmetic, logical)	Human factor	Computerized control

Source: constructed by the author.

Social studies in Internet have social and demographic characteristics. Studies of social and demographic characteristics of social network members are meant to reveal real and potential consumers of goods, services and ideas, in order to improve communications between the ones who offer these values and the target audience [5]. This can be done in

Internet, because many members have to fill the questionnaires when accessing some kinds of networks, offered by these networks. But some problems with reliability of data can occur. When building the nation-wide sample of the adult population for tête-à-tête interviews, we know how many people are in this population, how many people have primary, secondary or higher education; how many women and how many men etc., and it lays the firm ground for further statistical analysis. The problem occurs when interviews are conducted in Internet, because we do not have complete and reliable data on gender and age structure of Internet users (place of residence, age, education, gender and so on).

It is, therefore, necessary to obtain some data that can be used as the starting point or the reference point. For large research companies, this reference point tends to be the national representative sample of people who will be engaged as respondents in ad hoc studies conducted in Internet. These samples rely on statistical data, and they are carefully balanced. This type of studies includes the questions about Internet use, membership on specific social networks, percentage of persons accessing Internet from mobile devices, frequency of visiting certain websites etc. The data obtained in this way are to be used for correcting the sample in Internet, just like the census data is used to correct the average statistical number of the country's population, public opinion etc.

This approach allows for obtaining reliable social and demographic characteristics of social network members, and these studies are considered as representative in spite of their being performed in Internet. For example, Facebook is the largest social network in the world, containing great scopes of information about users; it can be regarded as an analogue of the so called demographic marking in conventional studies: date and place of birth, place of residence; user profession, including "history", or a possibility to write a short resume; information about education, including locations where various education diplomas were obtained; information about family, family status and relations; languages spoken by this person; religious beliefs and belonging; political views and opinions; address, phone numbers, etc. [10].

However, some users do not fill all the data, some of them limit their accessibility, but their percentage can be derived using data from conventional interviews.

Internet as a whole and social networks in particular offer an infinite source for very valuable information for companies and various organizations. In countries with high percentage of Internet users, the need for traditional surveys of public opinion and market research based on hard field works and direct contacts with respondents either tête-à-tête or by phone has actually disappeared [6].

In view of the above and bearing in mind that some networks have hundreds of millions of users, it can be assumed that a great deal of useful information can be obtained for various social, sociological or marketing studies and business analyses through Internet interviews.

Like conventional studies, it can be predominantly quantitative (based on measuring the frequency of a phenomenon or a correlation) or qualitative one, by measuring the traffic to some Internet pages and activity on these pages, or by analyzing attitudes or opinions of users of a social network, etc.

Internet searches and some Internet providers usually propose high quality and high precision tools measuring the number of visits of various websites. The most famous one is Google Analytics [7]. This and other similar decisions enable for something unimaginable in the conventional marketing (i. e. advertizing in conventional mass media): online monitoring of the campaign performance and efficiency, and the sales resulting from this campaign. There are two very important advantages over the conventional business and marketing:

- 1) quick and east control of the sales performance;
- 2) immediate taking of measures to eliminate problems and continuing improvements of promotions and sales.

Each actor of the market is, therefore, capable to monitor online how many visitors came to the website, their geo-location, the content looked at most or least frequently; duration of the website visit by website page; purchase time (if it is done in Internet) or where exactly (at which phase) they refused from the purchase.

Although the above said is more applicable for marketing studies, the main idea of using Internet as the platform for performing social statistical studies does exist.

Ukraine will witness the first online population census in 2020 using data of mobile operators and social studies instead of the conventional All-Ukrainian census in form of personal interview. The census will be performed by the following algorithm:

- collect anonymous data from mobile operators on the number of users and their location;
- collect data from the registers of children (offices for registration of the acts of civil status) and pensioners (the Pension Fund);
- make “calibrating” sociological interviews and adjustments;
- aggregate, sum up and translate the data into the mathematical model.

We are not experts capable of evaluating the Government’s competency with respect to the idea of online population census. However, we are fully aware that a new approach in statistical studies has to be based on the statistical methodology containing clear specifications of all the methodological and organizational processes, especially since it refers to such a mass-scale observation as the population census.

In view of the above said, statistical studies in social networks require a special methodology which content and form has to consider for all the peculiarities of Internet environment as a whole and its users in particular.

It is obvious that social networks, on account of their openness and information diversity, create the background necessary for the continuing improvement of social studies. Being conducted in the real time, they enable for a deeper grasp into the social world.

But one important question arises: to what extent this enhanced account of facts on human and society and this more robust research and analytical tools are capable to improve the quality of scientific perception and understanding of the society? The permanent engagement in fact accounts and scrutiny of great scopes of data can often hinder the desired view “from the above” or even distort the real process in space, in time, and logical sense. But this problem goes beyond the limits of this work.

References

1. Zbrytska T., Tabanova A. (2018). Aktualnist vykorystannia instrumentiv sotsialnykh merezh v osvithomu protsesi [The importance of using social network tools in the education process]. *Visnyk sotsialno-ekonomichnykh doslidzhen – Socio-Economic Research Bulletin*, 3(67), 117–130 [in Ukrainian].
2. Zmysna R., Fedushko S. (2016). Sotsialni merezhi yak forma suchasnoi komunikatsii: pliusy i minusy [Social networks as a form for the modern communications: pro and contra]. *Aktualni problemy humanitarnykh ta pryrodnychuykh nauk – Vital problems of humanities and natural sciences*. (pp. 149-151). Kyiv [in Ukrainian].
3. Dolzhenkova K. (2017). Stratehii ta instrumenty dosiahnennia populiarnosti ZMI v sotsialnykh merezhakh [Strategies and tools for studying the popularity of mass media in social networks]. *Zhurnalistyka: teoriia, istoriia – Journalism: theory, history*, 7, 28–37 [in Ukrainian].
4. Cattelan M., Firth D., Varin C. (2016). Statistical modeling of citation exchange between statistics journals. *Journal of the Royal Statistical Society: Series A*, 179(1), 1–63.
5. Durante D., Dunson B. (2016). Locally adaptive dynamic networks. *Annals of Applied Statistics*, 10(4), 2203–2232.
6. Glucksman M. (2017). The Rise of Social Media Influencer Marketing on Lifestyle Branding: A Case Study of Lucie Fink. *Elon Journal of Undergraduate Research in Communications*, 8(2), 77–87.
7. Google Analytics. Retrieved from <https://analytics.google.com/analytics/web/provision/#/provision> (date of access: 28.12.2019).
8. ITU: Committed to connecting the world. Retrieved from <https://www.itu.int/en/ITU-D/Statistics/Pages/default.aspx> (date of access: 28.12.2019).
9. O’Malley A., Onnela J. (2014). Topics in social network analysis and network science. ArXiv e-prints, 1–12.
10. Merchant G. (2012). Unravelling the social network: theory and research. *Learning, Media and Technology*, vol. 37, no. 1, 4–19.
11. Psychology of social networks. Retrieved from <https://www.seonews.ru/analytics/psikhologiya-sotsialnykh-setey/> (date of access: 30.12.2019).

12. Ofitsiynyi sait Derzhavnoi sluzhby statystyky Ukrainy [The official website of the State Statistics Service of Ukraine]. Retrieved from <http://ukrstat.gov.ua> (date of access: 02.01.2020) [in Ukrainian].

Bibliographic description for quotiong:

Holubova H. (2019). Using Internet Capacities for Social Statistical Surveys. *Bulletin of National Academy of Statistics, Accounting and Audit – Naukovyy visnyk Natsionalnoi akademiyi statystyky, obliku ta audytu*, 4, 7–13 [in English]. doi: 10.31767/nasoa.4.2019.01.