

MISUSE OF PERSONAL DATA IN PUBLIC OPINION POLLS – NEW EXAMPLES IN THE FORM OF INTERNET OF THINGS DEVICES AND APPLICATIONS

Vera Arežina, PhD¹

Faculty of Political Sciences, University of Belgrade, Serbia

Nenad Spasojević, MSc²

Faculty of Political Sciences, University of Belgrade, Serbia

Aleksandar Peković³

Abstract: This paper will analyze the importance of quantitative and qualitative research, but also their impact on the security of citizens who participate in them. Although they are of great importance both for science and for commercial purposes, personal data collected through research can be misused. During the pandemic caused by the COVID-19 virus, there was a rapid growth of online research, due to restrictions on movement and gatherings. This method of data collection is an example of how additional interests can motivate researchers to misuse personal data of respondents in different ways.

Besides, much bigger problem can be created by Internet of Things applications and devices that collect and analyze data using their software. Therefore, this paper will also offer solutions to limit and prevent potential misuse of personal data of users of these applications and devices.

Keywords: *analyze data, public opinion polls, data misuse, internet security, Internet of Things devices and applications.*

IMPORTANCE OF QUANTITATIVE AND QUALITATIVE METHODS FOR SCIENTIFIC AND PRIVATE SECTOR RESEARCH

In the social sciences, it is of great importance to apply various research methods in order to substantiate and expand existing theories or refute them with new knowledge. The application of qualitative, quantitative, and in recent decades mixed methods has enabled scientists and researchers to make an additional contribution to scientific reflections and answers to numerous questions. If we want to see



their importance for the study of socio-political changes, we can take the example of the elections in Great Britain at the end of the last century. The British Election Surveys (BES) made extensive use of quantitative methods, which was not enough to show the deeper causes of Labor's growth and their victory in the 1997 elections. By applying qualitative methods, some researchers have been able to show the deeper effects of this change, which only shows the importance of research methods (Devine, 2005).

Regarding the use of methods, there were great differences in the views of many authors and researchers. While some applied mostly qualitative methods, others firmly adhered to quantitative methods. The main reasons are their ontological and epistemological determinations. For more than a decade, mixed methods have been used to overcome this division between qualitative and quantitative methods. Mixed methods are the best for achieving new knowledge in order to complement the shortcomings that exist between them. In this regard, very famous methodologist Professor John W. Creswell explains the importance of the application of mixed methods, as well as the models for the application (Creswell 2014).

Mix methods also have great importance for any research within the private sector. Namely, it has always been of great importance to get information about what consumers of products and services think about them, as well as what drives them and influences their decisions when choosing. For these reasons, surveys of consumer attitudes and preferences became more common during the second half of the twentieth century. It is also a period of extremely great rise in the development of the economy of many countries, as well as their mutual trade.

All large companies have people in their teams who deal with marketing and market analysis, in order to gain customers. According to Kotler and Keller's analysis, the process of making decisions about buying a product consists of five phases, which are: understanding the problem, seeking information, evaluating alternatives, buying decisions, and post-purchase behavior (Kotler & Keller, 2006). Marketing teams are obliged to follow these phases, the changes of which can lead to a sudden withdrawal of customers from the products or services of a certain company. For such reasons, public opinion polls with citizens are important in order to examine their preferences and the reasons that influence to their choices for certain products and services. It is very interesting that with the development of Internet programs and applications, there has been an increase in sales over the Internet. Although many criticize this type of trade because it goes beyond the framework of legal regulations or is a matter of tax evasion, it is an increasingly common form of purchase, especially among the younger population. This certainly represents an additional challenge for companies, and forces them to start thinking about additional offers and checks whether such types of sales would suit the buyers of their products (Škare, 2006).

PROBLEMS IN CONDUCTING PUBLIC OPINION POLLS AND WAYS OF MISUSING PERSONAL DATA OF RESPONDENTS

As we have previously stated, there are quantitative, qualitative and mixed methods for conducting public opinion research. A large number of researches are conducted daily for the needs of science, politics or commercial purposes and many other reasons that strive to find out the attitudes of citizens on numerous issues. However, in order for a research to be conducted and its results presented to the public, it is necessary to prepare a lot of things, i.e. to start the analysis and creation of the planned research project before starting the data collection. When noticing a certain problem, it is necessary to first analyze what the current scientific and social fund says as well as previous research on that topic,



if any. When choosing a research topic, it is necessary to have a research project. It includes a draft of a scientific idea, various plans, as well as instruments for conducting research. Certainly, the draft of a scientific idea is the most complex, and at the same time among the most important elements, because it formulates the problem to be researched, determines hypotheses, indicators and the subject of research, analyzes previous research. In addition, it presents the scientific and social justification, as well as the goals that this research seeks to achieve (Marczyk, DeMatteo, and Festinger 2005, 16; Milosavljevic & Radosavljevic, 2003).

When it comes to a concrete example of the application of research methodology in scientific papers, it has become more present. For example, master's theses more often, in addition to description, explanation and classification as the most common scientific goals, also contain research. Empirical research that students in Serbia apply in their master's theses is useful because they give additional importance to science, and the topics they cover do not only have a descriptive element, which is usually the most present when we talk about these works (Arezina, 2021). An example of this is the master's thesis "Air pollution in the Republic of Serbia as a consequence of energy efficiency", in which the author Nenad Spasojevic applied the results from his quantitative public opinion survey, gathering the data using the mixing technique of online and telephone research. In that way, in difficult conditions caused by the COVID-19 virus, because of which the realization of the research was specific, as the author states, the chosen techniques were used to obtain data from respondents of different groups and territorial distribution, which were defined before the research (Spasojevic, 2020).

Besides, doctoral dissertation represents the original and independent scientific research and therefore the highest level of final works, with achieving knowledge in specific scientific field. Consequently, many faculties and universities have a number of regulations and guidelines that guide researchers and students to make the greatest possible contribution to their doctorates (Arezina, 2021).

While in research for the needs of science, researchers are expected to come up with new knowledge that will encourage further development of science, in practice there is also research that someone orders. These surveys must also follow the above steps in order for the data to be obtained to be adequate. Unlike scientific research, examples with commissioned research are different because we have the wishes of a particular client, and it often happens that they are more complex to implement. This refers to situations with the necessity of contract existence, frequent ignorance about the subject of research by the client, and possible disagreements and misunderstandings. The latter are a consequence of the fact that the client expects to receive practical proposals for solving his problem, while the researcher on the other hand comes to the origin of the problem and how to establish the basis for solving the problem, but also strives for opportunities to provide practical and concrete solutions (Gacinovic, 2017).

Researches are facing with several problems in the realization of public opinion polls. For the purposes of conducting public opinion research, it is necessary to have several important things. First, the professional knowledge and skills of the researcher are needed, then to have the experience gained through practice in conducting research. In addition to this, it is necessary to have human resources, because on the example of field research, it is impossible for the entire research process to be conducted by one person. Important element are the financial and other resources and capacities that would be desirable for the researcher or an agency to have. However, it often happens that the researchers themselves, disregarding the previous examples of the constituent elements of the research vocation, find themselves in a situation where they cannot realize the research as originally planned, or rather what it should actually look like in a professional framework. Consequently, there are situations where researchers without experience, originality, competence (certified knowledge), methodological cul-



ture, but also other shortcomings, bring the realization of research into question. This actually means that the data obtained by researchers are of questionable relevance, especially when it comes to the way they were obtained and who the respondents were according to the sample of researchers (Adamovic, Ivic & Vukovic, 2017).

Determining the sample is one of the most common problems in public opinion research. When starting the research, it is necessary to single out a certain unit from the entire observed mass, which would be studied, i.e. which would be a research sample. In his book "Statistical Techniques and Procedures in Political Science Research", Dževad Termiz, in the section on statistical mass and sample in political science research, defines these terms both in theoretical and practical terms, applying natural sciences as well. In this way, he also gives formulas on how to determine the size and validity of the sample itself (Termiz, 2006). Many researchers have difficulty to determine the pattern in their research, which is a consequence of either neglecting this element as an essential or more often non-existent knowledge of how it is determined.

The mentioned sampling is an increasingly common problem in Serbia. Numerous researches appear in the media that do not have a good methodological basis to be valid. In addition, research that has been unsigned, without data who conducted it, when and where, based on which sample, has been a huge problem in Serbia for years. Thus, we come to the conclusion that in Serbia, first of all, the most frequent, quantitative public opinion polls are largely inadequate, i.e. that they give a distorted picture of reality in terms of citizens' attitudes (Pavlovic, 2021).

Although mix methods have been applied more than decade, it is necessary to point out some of the qualitative and quantitative methods. Qualitative methods include observation, case studies, interviews and focus groups (Arežina, 2021). When it comes to quantitative research, in addition to scales and experiments, the survey is the most common (Creswell, 2014; Arežina, 2018). In the field of surveys, telephone (CATI), field (CAWI) and in recent months increasingly frequent online surveys stand out. For the purposes of this paper, examples will be presented with the problems of abuse and omissions during the conduct of research, and above all with surveys, given that unacceptable actions occur much more often.

According to Moser, "a survey is a technical procedure for collecting facts of material by combining a statistical method of sampling by interview or questionnaire" (Moser, 1962). According to this definition, we are actually talking about a data collection technique better known as testing. When it comes to the quantitative method, it has evolved over time, and so through the field and telephone, the survey became possible through mail survey options by sending e-mails, but also through other modern ways (Djordjevic, 2012). Today, online surveys are also widely used, which are the simplest and most cost-effective way for many researchers to collect data from respondents.

When collecting data for research purposes, data misuse is also common, which can best be seen in the example of online surveys. Thus, when filling out an online questionnaire (take the most common Google drive questionnaires as an example), researchers finally give respondents the opportunity to leave their e-mail address to inform them most often about the results of that research, and often about future research. However, as it is difficult for some researchers to reach respondents, they use this option to create their own database of respondents, by collecting email addresses. With each subsequent survey, they reach out to send emails to previous participants. In this way, the average user of Internet services is exposed to spam, by using his e-mail address for other purposes. As the COVID-19 pandemic spread, it also led to a huge prevalence of online surveys, and this led to further development of this problem (Google surveys, 2021). In this case, the e-mail addresses of the respondents are misused for other purposes that were not originally planned and presented to the respondents.



Figure 1 Example of an online questionnaire and questions at the very end, where the respondent is given the opportunity to leave his / her e-mail address in order to get the results of the research

In some field research in our country, there are two problems that threaten the privacy of respondents' data, and which can be used by researchers as in the case of online surveys to create a database. However, specific problem here is the part of the questionnaire called the header, in which data are entered only by the respondent (street and building number, floor, apartment number, and if possible name and surname). In addition to these data, the respondent is required to leave contact telephone number, in order to be contacted by the researcher to check whether the questionnaire was conducted. In addition to this, another problem is the use of GPRS navigation, which monitors the movement of interviewers. It is used to prevent interviewers from completing surveys outside the defined area of research. By using these applications, the privacy of the respondents can be endangered, because the place of residence of the respondents is found out, which can be interesting for researchers during the next research. In this way, there is a completely questionable anonymity of the research, which is guaranteed to the respondents before the very beginning of the research.

Figure 2 Example of a field research questionnaire, with a header, a couple of initial questions about the respondent and a brief description of the research

In the case of telephone research, we have situations that are similar to the problems from the previous two cases. The problem in telephone research is the way in which the researcher comes to the telephone numbers of the persons he would like to interview. In Serbia, for years, numerous agencies and companies have large databases of mobile and landline numbers of Serbian citizens, which leads to the conclusion that some mobile operators do not respect the privacy rights of their users and the data becomes publicly available to anyone.

These examples represent situations in which it is possible for the personal data of the respondents to be misused by the researcher during the research. It is important to note that these examples are not a constant occurrence and that not all researchers make these illegal ways to obtain certain data through the misuse of personal data of citizens. However, the actions of the citizens themselves are also a big problem. Namely, by using internet content and various applications, citizens often unknowingly provide their personal data without caring about their security. Besides, these data are important to large companies (Suceska & Hanic, 2014). It often happens that companies or smaller companies use the data collected from their former customers or service users in order examining their attitudes. This is a trend of checking consumer habits, but the problem is how data can also be used by the private sector.

On account of the mentioned unsafe access by citizens to the use of Internet content, it is necessary to point out that there are several omissions in the use of Internet services and applications. Thus, citizens are insufficiently aware of the potential dangers, even when they know about the risk, they are not too interested in preventing it. Also, they do not read too extensive terms of use and accept them lightly, but they also agree to share their personal data opposite to what they get for use. The results of numerous researches on the use of social networks and internet content show very interesting data. The reason for that is that they show that e.g. in the USA, despite the large number of cases of personal data breaches, citizens still do not pay enough attention to the protection of their data on the Internet (Mitrović, 2020).

DATA COLLECTION VIA INTERNET OF THINGS DEVICES AND APPLICATIONS AND WAYS OF THEIR MISUSE

Almost all companies engaged in the production of modern technologies in their business use comprehensively collected data in order to tailor their product to the needs of their customers. As manufacturers strive to make the use of a new product intuitive, the necessity of using the results of different types of research becomes inevitable. By using this data in different stages of product development, starting from the appearance of the product, through its functions and capabilities, to the marketing strategy itself, companies believe that using these data will better position themselves on the market. However, after collecting them, as we will see, almost nothing stops these companies from selling the collected data to third parties. Doing so directly violates the privacy of users, as well as ethical principles. The question is: who is the owner of the data we leave behind when using information and communication technologies?

To begin with, we will briefly look at how companies collect user data. First of all, we should start from what data are important for the further development of products and services, i.e. what are the data collected (Freedman, 2020):

- Personal data - which includes data such as ID number, browser cookies, device ID number and others;



- Usage data - talk about how users use the website, applications, social networking sites, paid advertisements and other content;
- Behavioral data - this group includes data on how the user uses the device, its purchase history and other quantitative data;
- Attitudinal data - they consist of measurable indicators of customer satisfaction, product demand, etc.

The methods by which companies obtain these data are: by direct inquiry of users, indirect monitoring of users or obtaining information from other legal entities (Freedman, 2020). Consequently, the question arises why do companies act like this? The answer is simple - the goal is to increase profits. Companies achieve this in two ways: by using the collected data to continue to develop their products and services and by selling the collected data to third parties. It is this sale that has led to the creation of new business models, where companies are established whose business concept is precisely the trade in this data - data brokers. They collect the data, sort it and process it, after which their results are ready for sale, and their clients are companies that have opted for the method of so-called personalized marketing. All this happens without the knowledge of the device users. Interesting data from the user privacy survey, conducted by Cisco Systems, states that only 32% of citizens are active in protecting their data, and 57% of respondents believe that they cannot adequately protect their data (Cisco Systems, 2019).

It can be seen from the above that this whole area of collecting, processing and selling personal data of users is a social problem, and looking at it, adequate legal regulations must be established for it. For now, legislation in this area around the world is rather scarce. Conditionally, the strictest law that directly concerns the protection of users' personal data is the California Consumer Privacy Act.⁴ In short, this law regulates the right of users to know what data is collected about it, how they are shared and used, then the right to delete collected data as well as the possibility of refusing to sell their data (California Consumer Privacy Act, 2018).

Although the previous part of the paper refers to the general principles of collecting personal data through smart devices and services (whether they are smartphones, personal computers, browsers, etc.), we should not forget that modern technologies such as the Internet of Things belong to this group, and that it is thus part of a spectrum of devices that collect the personal data of their users. For this reason, this general methodology for collecting personal data on smart devices also applies to the Internet of Things. As for protection against this type of information outflow, the principles are also similar, but the Internet of Things has the added problem that these devices usually do not have enough processing power, and if the manufacturer did not take security into account when designing the device, there is little possibility of installing any subsequent security programs (Jing, Vasilakos, Wan, Lu, & Qui, 2014). Other limiting factors of production do not support this, such as energy consumption, memory size, necessary time for production (Jing, Vasilakos, Wan, Lu, & Qui, 2014).

A good example of understanding how the Internet of Things collects personal data is the fact that these smart devices often collect location data and query data (Jing, Vasilakos, Wan, Lu, & Qui, 2014). Although this problem occurs at the application level of the Internet of Things, they certainly form a unique system, and for that reason this problem cannot be related to the device on which the application itself is located (example: application on a smartphone). The analysis of these data can lead to the residence of the user, their financial situation, behavior, health condition, etc. which the user unknowingly left behind during the search, e.g. for the nearest restaurant (Jing, Vasilakos, Wan, Lu, & Qui, 2014). It is a frightening fact that, in addition to the small number of users who even know that

4 California Consumer Privacy Act of 2018 [1798.100 - 1798.199.100] available at: https://leginfo.ca.gov/faces/codes_displayText.xhtml?division=3.&part=4.&lawCode=CIV&title=1.81.5



such data are collected, there is little they can do to protect their personal data. They are faced with the choice between using the device and providing their personal data, and where without the consent of the company to dispose of the user's data - the user is prevented from using the device which he owns.

Although all this may seem as if it did not happen so often and the consequences are felt only by individuals paranoid about their privacy, in reality the opposite is true, as evidenced by scandals where Internet of Things collected users' personal data. Thus, the smartwatch called TicTocTrack, intended for child safety in its design, had security flaws that allowed others to easily access the GPS data of the watch, as well as for the attacker to present himself as a "parent" and thus collect personal data of the child (Kirk, 2019). An additional problem is the fact that due to a software error, by accessing one account, it was possible to access all other accounts registered in the territory of Australia (Kirk, 2019).

These devices do not have to directly possess data of importance, but they can be an entrance to the network itself, which is a kind of problem not only for the individual, but also for companies. Although this type of data does not belong to personal data and thus is not the topic of this paper, it is important to be mentioned, given that they can have significant material consequences for the work of the legal entity, and thus indirectly on the micro and macro economy in which that company is located (primarily depending on the size and contribution of the legal entity). An example of a device that is found in the homes of users, but also in offices, are smart printers, which make up 18% of the Internet of Things, and as much as 24% of total security vulnerabilities (Unit 42, 2020).

CONCLUSION

In the modern world, collecting data on users is a very important element of the business of many actors. In this regard, these data are of increasing importance for both the private sector and science, all with the aim of easier access to the views and opinions of citizens and respondents. For this reason, the question of the moral and security conduct of researchers, agencies and companies in the implementation of these public opinion polls is increasingly being raised.

Taking all the above into account, the question arises: how to stop and dispute this problem of collecting and using personal data by third parties (often without the knowledge and consent of the device user)? Also, the next question that arises is: who should stop it? If the responsibility is on the individual, the sad fact is that they can do little. It has already been stated that an individual is faced with the decision to either not use the device and thus protect their privacy, or to use the device but thereby "disclose" their personal data. For this reason, we believe that at the level above the individual, primarily at the state level and by law, the user of the device must be protected and the area of data collection by the manufacturer must be regulated. The first step in this solution is for the state, or an international body, to recognize the collection of personal data as a social problem. Only after that can we start finding the best way for producers to obtain objectively necessary data for their work, while keeping individuals safe.

Another element for solving this problem should be the education and information of users. Just as the state should recognize as a social problem the misuse of collected personal data, so the individual should understand that each of their interactions with the Internet of Things leaves behind a trace that someone can use. Security culture is a key. Also, when citizens are informed about this problem, using democratic mechanisms they can influence decision makers and insist that this area be regulated. By establishing such a system, society can prepare for this type of new technology, the number of which will certainly not decrease in the future.



REFERENCES

1. Adamović, Ž, Ivić, M, & Vukvović, V. (2017). *Metodologija i tehnologija izrade naučnih radova*. Banja Luka: Univerzitet za poslovni inženjering i menadžment Banja Luka.
2. Arežina, V. (2021). Metodologija istraživanja u političkim naukama. *Srpska politička misao*. XXVIII, vol. 71, 273-292.
3. Arežina, V. (2018). Research Design in Methodology of Political Science. *Proceedings of ADVED 2018 – 4th International Conference on Advances in Education and Social Sciences*. Istanbul: ADVED.
4. California Consumer Privacy Act , 1798.100 - 1798.199.100 (State of California 2018).
5. Cisco Systems. (2019). *Consumer Privacy Survey*. San Jose: Cisco Systems.
6. Creswell, John W. (2014). *Research design: qualitative, quantitative, and mixed methods approaches*. Los Angeles: SAGE. Fourth edition.
7. Freedman, M. (2020, Jun 17). *How Businesses Are Collecting Data (And What They're Doing With It)*. Retrieved from Business News Daily: <https://www.businessnewsdaily.com/10625-businesses-collecting-data.html>.
8. Gaćinović, R. (2017). Prikupljanje podataka u procesu naučnog istraživanja. *Politička revija*. XVI, vol. 52, 137-156.
9. Devine, F. (2005). Kvalitativne metode. In: D. Marsh & G. Stoker (Ed.), *Teorije i metode političke znanosti*, Zagreb, Hrvatska: Fakultet političkih znanosti.
10. Đorđević, M. (2012). *Istraživanje politike: Bihevioralni pristup*. Belgrade: Službeni glasnik.
11. Jing, Q., Vasilakos, A. V., Wan, J., Lu, J., & Qui, D. (2014). *Security of the Internet of Things: perspectives and challenges*. New York: Springer Science+Business Media.
12. Kirk, J. (2019, April 15). *Australian Child-Tracking Smartwatch Vulnerable to Hackers*. Retrieved from Bank Info Security: <https://www.bankinfosecurity.com/australian-child-tracking-smartwatch-vulnerable-to-hackers-a-12376>.
13. Kotler, P. & Keller, K. J. (2006). *Marketing menadžment*, Belgrade: Data status.
14. Marczyk Geogrey, DeMatteo David, and David Festinger. (2005). *Essentials of Research Design and Methodology*. New Jersey: John Wiley & Sons, Inc. p. 29.
15. Mitrović, M. (2020). Sloboda izražavanja i zaštita podataka o ličnosti na internetu: Perspektiva internet korisnika u Srbiji. *CM : Communication and Media*, XV(47) 5–34.
16. Milosavljević, S, & Radosavljević I. (2003). *Osnovi metodologije političkih nauka*. Belgrade: Službeni glasnik.
17. Mozer, C. A. (1962). *Metodi anketiranja u istraživanju društvenih pojava*. Belgrade: Kultura.
18. Pavlovic, N. (2021). *Metodološki problemi u sprovođenju kvantitativnih istraživanja posredstvom interneta u sociologiji*. Niš: Univerzitet u Nišu – Filozofski fakultet.
19. Spasojevic, N. (2020). *Aerozagadenje u Republici Srbiji kao posledica energetske efikasnosti*. Belgrade: Fakultet političkih nauka, Univerziteta u Beogradu.
20. Starr, D. (2020, Avgust 10). *Here's Why You Should Be Concerned About Cameras in Robot Vacuums*. Retrieved from Safety: <https://www.safety.com/heres-why-you-should-be-concerned-about-cameras-in-robot-vacuums>.



21. Sućeska, M, & Hanić A. (2014). *Zaštita ličnih podataka u BiH i Srbiji. Sinteza - Međunarodna naučna konferencija Univerziteta Singidunim*. Belgrade: Univerzitet Singidunum.
22. Škare, V. (2006). Internet kao novi kanal komunikacije, prodaje i distribucije za segment mladih potrošača. *CROMAR (Hrvatska zajednica udruga za marketing) i Ekonomski fakultet Zagreb*, 18 (1–2) 29–40.
23. Termiz, Dž. (2006). *Statističke tehnike i postupci u politikološkim istraživanjima*. Lukavac: NIK Grafit.
24. Unit 42. (2020). *2020 Unit 42 IoT Threat Report*. Santa Clara: Palo Alto Networks.
25. Wiens, K., & Chamberlain, E. (2018,, septembar 19). *John Deere Just Swindled Farmers out of Their Right to Repair*. Retrieved from Wired: <https://www.wired.com/story/john-deere-farmers-right-to-repair/>