USING DIGITAL FOOTPRINTS IN SOCIAL RESEARCH: 
AN INTERDISCIPLINARY APPROACH

Abstract

The study aims to clarify the concept of the digital footprint in jurisprudence and social sciences and determine its meaning for interdisciplinary research. The subject of this work is the analysis of the concepts of “digital footprint”, “digital reputation”, “digital image of a person”. The article notes that special machine techniques are required to process digital traces. The ways of storage, transmission and use of digital traces are considered. The research identified the main problems of using digital information in public relations. Digital traces can be used in various studies in the humanities, social sciences: sociology, law, economics, psychology using interdisciplinary research methods.

The result of the study was the conclusion that digital information will have a digital footprint only when it is transferred from one user to another in an online environment, uploaded to a social network, etc. (in the case when it is possible to “trace” its movement). Thus, the “digital footprint” is the ability of information recorded in digital form, to leave special “marks” on the route from subscriber to subscriber, the ability to track such marks, receive information about its movement and transformation in order to collect, process and analyze these data.

Keywords: digital footprint, digitalization of public relations, artificial intelligence, “digital image of a person”.

Introduction

Currently, the amount of digital information is growing rapidly. The impact of digital information on people’s lives should not be underestimated: the more information about people, their actions or some social phenomena appears on social networks, the more opportunities it can be used for research. This information leaves a digital footprint, the study of which can be used in sociology, psychology, law, economics and other social sciences. Digital information and its footprints can be used in the study of the demand for products, in marketing research; in psychology (for example, to determine the psychotype of a person or a group leader), in jurisprudence (to facilitate the process of investigating and collecting information about a crime), in sociology (for conducting various sociological surveys, building models, determining the level of unrest in society, etc.). The study of digital footprints and their use in the social sciences will solve the
problems of faster information processing, save time and effort, and will lead to more accurate research results.

The purpose of this article is to analyze complex interdisciplinary research in the context of identifying the possibilities of using digital footprints in various social sciences.

To achieve this goal, the following tasks have been solved: the concepts of “digital footprint”, “confidentiality in the network”, “digital reputation”, “digital image of a person” have been defined; methods of collecting, evaluating and analyzing digital footprints in sociology, psychology, jurisprudence, economics have been analyzed; studied the features of the circulation of digital information; analyzed the work of Russian and foreign scientists on digital information and digital footprints; suggested ways (including interdisciplinary) of using digital footprints in social science research.

It is the interdisciplinary approach that is new in the study of this issue since until now, and there has not been a comprehensive study of digital footprints from the position of several sciences at the same time. The results of the research allow to expand the area of using digital footprints and improve the quality of research.

Theoretical Framework

The use of computers has firmly entered our lives. It has become an integral part of the activities of not only individuals (games, correspondence, watching videos, searching for information), but also legal entities (electronic document flow within organizations, relations with banks and regulatory authorities, submitting reports, submitting applications), as well as public authorities (for example, the Russian portal “Government Services”). In most modern states, systems of “electronic government” (e-Government) are effectively operating, and attempts are being made to introduce “electronic justice” (e-Justice). In Russia, a full transition to the provision of significant state and municipal services in electronic format is planned from January 2023.

Thus, we are witnessing the growth of digital public relations, the transition of public relations to a digital environment and the digitalization of society as a whole.

The number of Internet users and the amount of information placed in it is growing, the quality of its use is changing - often up to 80% of information that used to have a material expression (photographs, books, videos, texts, etc.) is now stored on a computer, on a server or online disks (Yandex Disk, Google Disk) in digital form, without duplication in material form. The amount of information located in local networks, the World Wide Web, as well as in computers and other hardware (any technical devices on which information can be recorded - a smartphone, tablet, modem, etc.) is practically endless. The penetration of such information into all, without exception, spheres of human activity and society has led to the creation of new information categories: “virtual environment”, “electronic document management”, “digital footprint”. These categories are now becoming more widespread and require a separate study.

Information that once entered the global network remains there forever and looks like a digital footprint. Consequently, such information has almost limitless volumes and is big data (BigX Data), which requires special methods of assembly, evaluation and analysis.

Digital information is information encrypted with a binary code - 1 and 0. The specified

---

1 The instruction of the President of the Russian Federation V. V. Putin - https://tass.ru/obschestvo/9705537.
digital code forms a unique system that allows you to save (to record), transmit (to send from one hardware to another), and most importantly - recognize it when received by the addressee. Thus, this information comes to the user without transformation. Such information is digital by nature. In our opinion, in this case, the name “electronic information” is inapplicable (as some authors say about this type of information).

The term “electronic information” owes its origin to the storage medium, that is, the device with which digital information is transmitted, received and stored. The appearance of this term is associated with electronic computers, the appearance and popularization of which led to the penetration of digital information into everyday circulation. Thus, the legislator attributed any information encoded in the form of an electrical signal to computer information, although, as A. A. Fatyanov (2017) correctly notes, not all electrical signals are computer information. Confusion in the use of the terms “electronic”, “electrical” and “digital” leads not only to the substitution of concepts but also to their incorrect use in research.

The information that is generated, transmitted, stored on computers and other hardware is not just an electronic signal, but a signal that needs to be “digitized” so that it can be processed by a computer and become a digital signal. The encoded digital signal is constant, not subject to accidental or mechanical changes (experts detect deliberate changes), can be recorded both on electronic media (flashcards, disks) and on non-electronic media (optical disc) (Kalitin, 2014). Hence the conclusion: the information transmitted in the form of a digital signal is digital, and it is more correct to call it “information in digital form”.

Regardless of the content of the information itself (it is a photo, video or text), it remains digital until it is displayed on a screen or paper. This information is perceived by a person in a transformed, decoded (recoded) form.

Note that in foreign practice there is a clear separation of the terms “digital” and “electronic”. Understanding the differences between these concepts is important since research can use both digital information and electronic sources of digital (computer technology) and analogue information.

Any information (digital or analogue) acts as a source of data for research in various branches of science, including for interdisciplinary research. It was said above that the specificity of digital information lies in its machine processing, which allows not only to reduce the processing time but also to increase the volume of investigated data, which gives the advantage of digital information over analogue.

Currently, there is no clearly fixed, established opinion about the concept of “digital footprint”, although there are more and more studies on this topic (Boyarkina & Boyarkin, 2016; Glushkova, 2020; Tohtieva & Fulin, 2020). Most often, the digital footprint is presented as a unique set of actions on the network (Vangie Beal), or as a “digital shadow”, “electronic footprint”, “Internet footprint” (which is a web browsing trail (cookies) (Garfinkel & Cox, 2009) This term is usually applied to an individual user. However, it can also be applied to a legal entity.

Many digital footprints allow them to be classified. It is customary to distinguish two types of digital traces: active and passive (Girardin, Calabrese, Fiore, Ratti, & Blat, 2008). An active digital footprint appears as a result of the publication of any information by the user on the Internet, for example, on social networks. A passive digital footprint is formed when data is col-
lected about a user without their knowledge. Practice shows that users leave digital information intentionally or unknowingly. Persons interested in this information collect it actively or passively. A feature of digital information is its volume and shape. Since all this data is stored on the web, it is relatively easy to collect a lot of user data using simple search engines or using machine processing.

Thus, “digital footprint” is the ability of information recorded in digital form to leave special “marks” on the way from subscriber to a subscriber; the ability to track such marks, receive information about its movement and transformation in order to collect, process and analyze this data.

The study of digital footprints began a long time ago (almost from the moment of their appearance). In 2007, Tony Fish did much research in the field of marketing and revealed not only the prospects for their use but also identified the problems and warned about the dangers that this data can store in itself. In his work, he noted that the digital footprint could determine the needs and motives of a purchase, determine the purchasing power of a person, determine the reasons for the purchase, etc. (Fish, 2009).

In 2017, Katalin Feher conducted a study on online strategies for posting information by users on the Internet. As a result of the study, she concluded that the consequences of posting information on social networks can be unpredictable: only 70% of the information remains under the control of the user, and 30% of actions are not subject to control by the users themselves and can be used by unauthorized persons or intruders, including for the purpose of committing illegal acts (Feher, 2017).

Passive digital footprints can be stored in at least two ways:

- fixation in on-line databases: creation date; time and travel paths from user to user; The user’s IP address; file size, etc.
- placement in an off-line environment: for example, placement in files (as content), to which the administrator has access, who can determine what actions were performed with the file and when (for example, look at the log when the file was modified or moved). At the same time, the administrator cannot determine who exactly did this.

Active digital footprints can also be formed and stored in two ways:

- in the online space (when the user visits the site or his page on the social network where he is registered, publishes or posts some information);
- off-line (when the trace remains “inside the computer”);
- in a keylogger (fixing any changes in the clipboard (for example, when a user copies passwords, makes screenshots of information, or performs some other actions that at first glance leave no traces)).

Let us consider each of them.

1. “Confidentiality on the Internet” - it can be noted that it practically does not exist. A person needs to make much effort so that the actions carried out by him in the network are not interconnected. We have conducted an experiment on personal identification on the Internet using digital footprints. The study was conducted concerning two groups of subjects: group A and group B. In group A there were users who, being active users of the Internet and social networks, were registered there under their own names (nicknames in the social network corresponded to the surname and first name). In group B, users were also active users of the Internet and social networks. However, they were registered there un-
nder anonymized nicknames (nicknames in the social network did not contain an indication of their surname, contained a fictitious name or fictitious numbers). During the experiment, 70% of users from group B were identified until the surname was established. This became possible thanks to the use of methods of analysis, information evaluation, machine processing and the use of computer algorithms: a circle of communication (friends in a social network), mutual acquaintances, a city, an educational institution and/or place of work, a family were identified, and through the analysis of other people’s accounts it became possible to identify them.

As a result of this study, it was concluded that with the help of certain search programs and search algorithms, it is possible to identify almost any subject on the Internet. This proves that privacy on the network (virtual environment) is almost impossible, and it is almost impossible to apply the law on personal data to information on the network. Since the information is posted by users, it becomes publicly available.

2. “Digital reputation”. This category manifests itself through research using sociological methods (observation, survey/questionnaire, analysis of documents (websites and personal pages of social networks), modelling). The study analyzed the question of whom people trust in social networks; whether the reputation of the person who represents it influences the choice of the product (for example, advertising of bloggers). As a blogger, actresses, public people, TV presenters, show business stars and bloggers themselves (those people who work in the network: YouTube, social networks, etc.) were considered.

Our analysis showed that in 89% of cases, a purchase is made after the recommendation of a blogger whom the reader/user trusts, whom he believes, and has been subscribed to for at least 3 months. At the same time, each blogger has his own digital reputation, which is mediated by his actions on the network, statements, propaganda of any ideas, etc. This reputation can be both positive and negative, but unlike real life, it depends on the actions on the Internet. Digital reputation has its own digital footprint, which can show a change in the “level” of reputation. For the layman, “digital reputation” is important when solving some socially significant issues: for example, it may be needed (or “hurt”) by an applicant when entering a university; a person when applying for a job, when establishing personal contacts and relationships, etc. Any interested person (for example, an employer) can use social networks to determine the digital reputation of the applicant and draw some conclusions: how much time a person spends on the network, what is his manner of communication, attitude to projects, etc.

3. “Digital image of a person” is a kind of image that is created in social networks in relation to someone. A peculiarity is that this character can be both real (existing in life) and fictional (fake pages on social networks created to “cheat” subscriptions and votes). Particularly noteworthy is the image of a “digital person” who is identified with a real person. In this case, the methods of psychology are of interest. Using the methods of psychology, one can determine the psychotype of a person, predict his behaviour, and determine whether his digital image corresponds to reality. In less than 10% of cases, the “digital image of a person” and his real-life image did not coincide; in other cases, they were the same. The definition of the “digital image of a person” also occurs through the analysis of digital footprints.

The next step is to understand who can use the digital footprint and how. Persons using digi-
tal footprints are subjects of use. The subjects - users of digital footprints are marketers, sociologists, investigators (law enforcement officers), psychologists, financiers. Using digital footprints, you can determine the needs of people and predict the demand for goods of a certain category, the movement of people. At the same time, it is possible to carry out observation, both involved and not involved, to find out information that cannot be obtained offline (for example, the connection of subscribers via a cellular network or geolocation).

The digital footprint allows you to study the interests of people, determine the location of users; communities and groups in which this or that person is, his views on certain events, etc. (Vatlecov & Vatlecova, 2019). Digital footprints allow you to research communities and groups, make ratings and comparisons (for example, find out the rating of hospitals, schools, doctors, businesses, etc.). Moreover, all this information can be collected without the knowledge and permission of its owner. In order to prevent digital footprints from being trace, special measures must be taken.

Such a caveat does not mean that the digital footprint is a negative factor and that it needs to be eliminated. First of all, you need to understand that the formation of a “positive” digital footprint, on the contrary, avoids any problems, and can help in some issues.

A positive digital footprint can be created using the results of this study, and here we are already talking about a conscious online identity: emphasizing dignity, carefully monitoring what is published, etc. When studying the formation of a positive digital footprint, it is necessary to remember about the digital culture, which is only just appearing among users of the Internet space (Skivko, Korneeva, & Kolmykova, 2019).

Thus, the main actors in the use of digital footprints can be divided into three groups:

- official users - e.g. employers, government agencies;
- personal users who use digital footprints for personal purposes, as a rule, to clarify information about someone;
- business users, using digital footprints in marketing research, to study the level of sales, reviews, etc.

Let us consider each type.

1. Official users. Employers are increasingly using digital footprints. For example, in 2006, only 11% of employers checked the social networks of potential employees, and in 2017, more than 70%. Specialized companies (for example, Fama Technologies) appear that do such research on order for employers or other interested parties.

On the other hand, there are companies (for example, BrandYourself) that form a positive digital footprint for individuals and legal entities, various companies. Such companies analyze pages on social networks, comments, photos (including on the pages of friends) and determine what content gives a positive digital footprint and what needs to be removed from the page. Besides, such companies can create positive content, recommend some actions on the network, etc. Also, the user’s digital footprints can be considered by official state bodies (for example, when investigating crimes: identifying members of a criminal group, collecting and examining evidence expressed in digital form, etc.). In forensic science and criminal procedure, such studies have been carried out for a long time (Meshcheryakov, 2001; Lazareva, Olinder, & Perekrestov, 2019). However, it should be noted that interdisciplinary research on this topic is just beginning to be carried out.
2. Personal users using digital footprints for personal purposes. Studies show that 7 out of 10 young people, when they meet, always look at each other’s social networks. As a rule, these people lack any special skills; they use direct views, without the use of machine information processing methods and computer programs. By visiting the page of such a user (provided it is open), you can determine where the person was, whom you know, learn about his interests, level of wealth, behavioural habits, etc. Equally important is what other people write about someone on the Internet (these are passive digital footprints). They can also be found and analyzed without using special programs, but after spending a certain amount of time analyzing friends’ pages, reading comments under posts, studying hashtags, etc.

3. Business users (meaning users who use digital footprints in business). There are several directions here.

First, it is about digital reputation and brand presence on the net. Increasingly, companies, especially those focused on retail or providing services to individuals, are moving online. Among other things, this is facilitated by the situation with the Covid-19 pandemic, which arose in March-April 2020. Such companies are increasingly selling and promoting (marketing) their goods and services through social networks, and all these sales and type of account/site will constitute the digital image of the company (digital footprints).

Secondly, one cannot ignore online customer reviews (research has shown that 85% of customers are focused specifically on positive reviews, while about 65% of customers are ready to write and write these reviews).

Third, the presence of a business on the network becomes necessary, since already about 90% of users have bought something on the network at least once. More than half of the respondents begin their search for the desired product/service by analyzing offers on websites or social networks. A poorly designed website/account can simply scare off potential customers, so this area also needs to be paid attention.

Thus, it becomes obvious that digital footprints are an integral part of the life of any modern person, company, state. The absence of digital footprints is possible only if the subject is not online at all, and this is almost impossible during the period of digitalization of society. Thus, part of public services has already been fully transferred to the State Services portal (for example, registering a car, obtaining or replacing a driver’s license is impossible without registering on this portal). Therefore, in research, it is simply necessary to be able to use digital footprints to save time, improve the accuracy of results, and reach a larger audience.

Digital footprints in scientific research can be used in two ways:

- as a source of information for special, independent research in individual social sciences: in jurisprudence, psychology, sociology, economics, etc.;
- as a source of information for interdisciplinary research in several areas: for example, determining the type of personality (psychology) to determine its intended role (sociology) in a criminal group in the investigation of certain types of crimes (forensic science).

As part of independent research, work with digital footprints can be carried out in sociology, psychology, economics, jurisprudence, etc. It should be noted here that research using digital footprints and digital information will be carried out in the same way as when working with ana-
logue information. The same research methods can be used, answers to the same questions are obtained. The difference will be one - the advantage in the speed of information processing and its quantity. With the use of machine processing methods (which in this case will be only a method of digital information processing), one and the same array of information can be examined from different angles, additional connections between individual elements can be identified, cause-and-effect relationships can be determined, etc. The above were examples of exactly how digital footprints can be used.

With regard to interdisciplinary research, the emergence of digital information and the development of the doctrine of digital footprints makes a breakthrough in such research possible, increases their quantity and quality. This is precisely due to the use of special methods of machine processing and the use of special computer programs, which allows not only to process a large amount of information but also to combine information from different sources, different areas, which was previously problematic. Such studies can combine several areas at once: economics, sociology, psychology, jurisprudence, etc. For example, you need to study some economic indicator using the methods of sociology and make a decision based on the laws of psychology.

Research shows that companies, when collecting digital footprints of users, use only part of the information for the stated purposes (for example, to improve the quality of their work). Most of the information received is processed for “predictive” purposes, often without notifying the customer, and even more so the users themselves. Based on what the consumer is doing online, it is possible to predict an increase or decrease in demand, the emergence or fading of interests, a change in behavioural habits.

On the one hand, such tracking of digital footprints is capital and ubiquitous - it tracks all footprints that relate to a specific issue and get into the network. For this, machine information processing is used. On the other hand, the results of such tracking become a product (information base) that is in demand and can be sold on the information services market to interested parties (companies, political parties, public associations, etc.).

The undisputed leaders in the collection and analysis of digital footprints are Google and Facebook. Why is this information in demand and popular? If you have an accurate idea of the development of this or that phenomenon, then you can determine the period for a painless increase in prices for products, the launch of new types of services, etc. If this is a trading platform, then you can set individual prices for customers, focusing on their purchasing power - the same product/service can be sold at different prices: at a greater cost to those who can pay for it. This conclusion can be drawn from the analysis of the average check. Alternatively, having tracked the location and location of your customers, you can open an additional office or move a warehouse to a more convenient location, thereby reducing costs. You can also study the turnover of competitors and their prices, etc.

Thus, knowledge of the features of digital footprints, the order of their formation; peculiarities of search, collection, storage and analysis allows their wide use in interdisciplinary research.

Methodology

In the study, the results of which are presented in the article, both empirical methods and theoretical research methods were used. Empiri-
cal methods were used to collect data on such social phenomena as digital information and digital footprint, as well as data on their external manifestation. Using the methods of theoretical research, the collected material was processed, internal relationships were revealed (how the trace appears in the virtual environment, what properties allow us to speak of it as digital, etc.).

The following methods of empirical research were used in the study: observation (research of people’s behaviour in social networks), experiment (identification of a person on the Internet using digital footprints), comparison (research of digital footprints from the point of view of different social Sciences). Methods of theoretical research included: analysis (research of digital information itself and digital footprints, methods of obtaining it), modelling (building a model of the possible use of digital footprints in research).

Among the special research methods, we used sociological methods: systematization, survey, involved observation, and documentary method. Sociological methods are a way of collecting and processing digital information and digital footprints. They are the basis of interdisciplinary research, as they allow you to get data to answer questions posed by other social Sciences. In the study, using sociological methods, the following were carried out:
- registration of actions on the Internet—a survey, observation;
- data collection—a sample survey of accounts and sites;
- data processing and analysis—system analysis of sites, description of accounts and users.

We also used psychological methods for studying personality: analysis of data about the individual in the social network account, observation, comparison. Using psychological methods, we established cause-and-effect relationships, identified the motives of people’s behaviour, determined the prerequisites for determining the psychotype of a person, etc.

Special legal methods are necessary for analyzing the legal regulation of the issue of handling and protecting digital information and digital footprints. Among the legal methods, we would note the identification of conflicts of legislation in this area, a comparative legal method for processing the experience of legislative regulation of these issues in other countries.

Economic research methods are necessary since the collection, analysis and forecasts of digital information and digital footprints are necessary for the development of the economy or for obtaining information about economic processes. Thus, interdisciplinary research requires a comprehensive application of General and special scientific methods, including those from various social fundamental and applied sciences.

Results

The main result of the research is the identification of opportunities for applying an interdisciplinary approach to the study of such a phenomenon as the “digital footprint”. The results of the study allow us to expand the use of digital footprints and improve the quality of research.

The study showed that information that once got into the global network remains there forever, and has the form of a digital footprint, has unlimited volumes and is characterized by big data, which requires special methods and approaches to their research.

During the work, much attention was paid to such concepts as “digital footprint”, “online privacy”, “digital reputation”, “digital image of a
person”. It is noted that a digital footprint is the ability of information recorded in digital form to leave special “marks” on the way from subscriber to subscriber, the ability to track such marks, get information about its movement and transformation in order to collect, process and analyze this data. Digital footprints include two large groups: active digital footprints and passive digital footprints.

The results of the study showed that it is quite difficult to maintain confidentiality on the network as all data posted in open sources are open. To obtain this result, an experiment was carried out to identify a person on the Internet, namely in social networks.

Much attention in the study was paid to the concept of “digital reputation”. This category manifests itself through research using sociological methods (observation, survey/questionnaire, analysis of documents (websites and personal pages of social networks), modelling). We came to the conclusion that there is a positive “digital reputation” and a negative “digital reputation”, the formation of which is influenced by both the user himself and his environment, or special firms.

No less significant results were obtained when studying the category of “digital image of a person” - this is a certain image that was created in social networks in relation to a person. A special feature is that this character can be both real (existing in life) and fictional (fake pages on social networks created to “cheat” subscriptions and votes). The image of a “digital person” deserves special attention, which is identified with a real person. In this case, the methods of psychology are of interest.

It is highlighted that the subjects-users of digital footprints are: marketers, sociologists, investigators (law enforcement officers), psychologists, financiers. Three groups of user subjects have been identified: official users; personal users; business users.

The analysis of the ways of using digital footprints in interdisciplinary research is carried out.

Conclusion

Thus, we can conclude that digital information and digital footprints as part of it (manifestation) are becoming an integral part of society. No science can ignore their appearance and the opportunities that they provide for research, namely: saving time and increasing the amount of research information.

Digital footprints in social science research can be used in two ways:

- as a source of information for special, independent research in certain social sciences: in jurisprudence, psychology, sociology, psychology, economics, etc.;
- as a source of information for interdisciplinary research in several areas: for example, determining the type of personality (psychology) to determine its intended role (sociology) in a criminal group in the investigation of certain types of crimes (forensic science).

References


