Stanislava Nerutė KLIGIENĖ
Vilnius University, Institute of Mathematics and Informatics
Akademijos 4, LT-08663 Vilnius, Lithuania

Received: August 2011

Abstract. The paper surveys the risks and benefits which the user faces in the networked environment and how those challenges can be overcome. The question is how to measure the potential or benefits of such a complex phenomenon – the collaborative cross-domains in social media. As one of the solutions we propose to consider this in the context of digital tools and the entities involved into cooperation-collaboration: core researches, engineers, developing information systems and tools, marketing technologists, users-consumers of services and products. The ways of collecting data and the measures for protecting privacy issues of the data collected online, as they were applied during the last two decades are overviewed in this paper. There is no universal law protecting the privacy of online users in the global world and hardly will it ever be. For a while, only the awareness of the users, the professional Codes of Ethics and fairness of firms involved into collaboration could help them avoid pitfalls hidden in social media. The summary table shows at a glance benefits and dangers met in social by its explorers and users.

Keywords: social media, networking, digital footprints, data privacy, professional ethics.

1. Introduction

Internet and social networks have become an integral part of our life. Two decades ago nobody predicted that the sights of any place of Earth would be visible and interactively controlled in computer networks – the object is approached and viewed on the desirable side – that it would be possible to discuss and exchange the contents, to communicate with people wherever they were at the moment, to search in large data arrays in a remote way and instantly get answer. A violent recent development of the internet and social networks has surpassed all expectations challenges of scandals because of Wiki Leaks document publicity in 2010, political events in Northern Africa in 2011 have demonstrated a huge power of information technology (IT). It is supposed that Facebook and other social networks were helpful in organizing protests even in such counties of Africa where the penetration of the internet is not so noticeable. At first, their governments tried to forbid social networks, but soon they have changed their mind since it is better to use the internet to achieve their own purposes than to forbid. On that occasion, the USA promised to support the internet development in non-democratic countries, while China took more drastic censorship measures on the internet. In our so spasmodic digitalized age theoretical models and IT applications now have to be critically reconsidered. The former information
recipients in the Web 2.0 environment become information conveyers ever more often. In the references (Ala-Mutka et al., 2009; Verdegem, 2011), social media are defined as open internet-based and user-friendly IT applications that establish new opportunities for participants: in creating the contents (Wiki, blogs, Flickr, Twitter); taking part in social networks (Facebook, LinkedIn), in sharing opinions, and looking for information (Google, Amazon). Really, this creates a collective brain, which has never occurred in the history of mankind. Soon there appeared some critical voices (Fish; Hassan and Thomas, 2006) about the danger of social media as information about the users in aggregated and then shared. The motto of the book “Digital Footprint” by Fish was: “A two-sided digital business model where your privacy will be someone else’s business!” In a networked society there emerge new powerful and interesting links that should be explored and considered. The aim of this survey is to discuss the benefit and threats of social media in the concrete context of users and their creators by destroying two myths contradicting each others: (i) techno-enthusiastic vision as to unlimited IT abilities; (ii) exaggeration of dangers and pitfalls hidden in the new technology. Let us elucidate WHEN and to WHOM the benefit may become a real danger and in what ways we can overcome imaginary or real threats.

Figure 1 illustrates interactions in the interdisciplinary activity spheres – theoretical and applied-involving users into the process. Data of users are the object of research, and the main aim of the whole process in meeting the needs of users (also taking into consideration the profit of service providers). It is of interest to note that the mathematical methods and principles, used in the new technologies, are almost the same as that of the past ages. Classification rules of multidimensional data, algorithms of automated teaching, statistical accuracy metrics, models of Bayes networks, Pearson’s correlation, accumulation forming and solution theory are but several methods and models. Well-known in the classical mathematics that currently are every widely used in creating modern IT tolls and applied programs for data mining and analysis. Let us consider the process that takes place in a networked society (Fig. 1) more in detail. Using mathematical methods, software is developed that is necessary to accumulate and analyze the users’ data, mar-

![Fig. 1. Major spheres of activity and their relationship, realized in social media of the internet.](image-url)
marketing methods and applied in producing services for users and a profitable business for enterprises. The data are collected, analyzed, and stored. The internet and social networks are the basis of the process where the data are collected, while the developers it ranked from the base of the pyramid, respectively are: scientist, researchers of theoretical – mathematical basics, principles and methods, engineers of IT who create systems, networks and applications, founders of social networks, managers, firms, that develop their business in this environment, and users whose data are collected and used for producing the necessary services (or sometimes just insistently offered).

As seen from Fig. 2, users in social networks are distinct in many aspects. From the viewpoint of participation, some of them are active creators; others are “digital tourists” that are looking round (exploring) in the digital medium; in line with their habit distinctions: some of them born in the digital era, accept IT tools as a natural endowment, others are the so-called “digital emigrants” who have to learn to conform to the digital technology. Users are also different according to other socio-economic factors: education, social status, dwelling-place, etc. Even more differences arise when classifying according to a user’s understanding about a digital footprint; some of them know and take care of collected and accumulated their data, while others have not got the slightest idea about data collection and their possible handing over to the companies having interest in them, still others are aware and think that it is worth giving data in exchange for better and innovative products, services; and, finally, some users are aware of that only partly, but underestimate the risk emerging due to a large amount of the data accumulated. So, we will consider the advantages and threats provided by social media in the context of the participants mentioned above. We are living in the age of information abundance, when the news is instantly announced on the internet and most distinct viewpoints are published on the internet or in traditional press. Therefore it is very easy for a reader or a participant of the new media, to be mistaken. Selection of information and capability to filter and objectively estimate it becomes the most important and indispensable property for a successful activity in the age of information.
In Second 2, we consider the principal notions of the digital era and how they are manifested in the present society. The Section 3 is meant for discussion of the challenges and problems of the digital era as well as for search of the ways of their solution in separate countries by making an attempt to overcome those e-challenges. The Section 4 overviews the whole complex and conclusions are drawn about the advantages and threats caused by social media for each participant of the process. For data analysis we use computer technique power that enables us to extract the necessary knowledge from large scale arrays. Researches of this kind seem to be on the verge of scientific ethics as, without any suspicion of an individual, the data about his/her behavior in the internet are aggregated and he is appointed to one or other group and products are offered to him that he had not asked for. Is that a smaller infringement of an individual’s rights than in biotechnological researches where the ethics is a traditional priority? The necessity of IT professional ethics, as one of more important ways of decreasing danger, is discussed in Section 5.

2. Notions of the Digital Era

Specific methods of information communications technologies, developed in line with the needs of the digital era, indicate a shift from social informatics to social intellect, achieved by investigating social dynamics, by modeling and analyzing social behavior, by crating artificial social agents and by generating and controlling social knowledge. That is how behavioral targeting comes into being that can operate only if the data are collected. Earlier the data were collected by applying various polls, census by talking representative statistical samples and drawing from those well-founded statistical conclusions about the population researched. That was a complicated and laborious process. After the digital interaction has become common place, data collection was simplified a great deal, however we can’t ignore arising juridical problems. We should essentially revise online collected data processing methods. People operating online leave the data about themselves via the so-called digital footprints and social graphs. Let us define these and other notions, necessary for a deeper understanding of a problem, more exactly.

Digital Footprint and Digital Shadow

This notion appeared first and was called “slug trail” (Negroponte, 1996), later Tim O’Reilly called it “data exhaust”. The notion was used then only to define the information that remained after browsing on the internet. At present this phenomenon is named as a digital footprint and it reflects the data created and used no matter by which digital devices. The digital footprint is a trace that remained after an individual’s actions in a digital environment, including the use of TV, mobile telephone, internet, sensor or any other device. The digital footprint connected with WWW is sometimes called as an internet footprint, a digital shadow or a cybernetic shadow, i.e., the information left after browsing on the internet, including that obtained via cookies. The term of footprint is applied not only to individuals, but also to firms and enterprises. A digital footprint is a combination of activities and behavior when the entity under consideration (a person or
something else) is acting in the digital environment. These may be log on or off records, address of visited web pages, open or developed files, e-mails or chat records. Digital footprints of this kind are accessible to data mining when the interested parts seek to learn more about some entity or individual. The footprints show the sites where we have been, how long and how often they were visited. That is as if the moments of memory. They do not include either data of personal identity or numbers of passport, bank and insurance accounts. However in some cases, those exhaustive data arrays allow us to reveal a person’s identity almost exactly. The digital footprints collected from mobile communications internet entries, interactive TV, as well as metadata of these actions show the truth about individual persons or enterprises, so that control of these data becomes a field of severe fight between the interested parts.

**Footprinting and Reality Mining**

Input into a digital footprint embraces a lot: day time, site, search results and keyword, created or reviewed contents, digital activity and data from sensors and social environment. Some of the data may arrive from depth of internet data IP which is called footprinting. Recommendations, contextual adaptation and personalization are created from such collected inputs and data analysis. A part of this analysis is called reality mining. The reality mining explores the human interaction based on wireless devices such as mobile telephones and GPS localization which yields a very exact image: what people are doing, where they are going and with whom they are communicating. That is even more trustworthy than stories of the very individual. Reality mining is one of the important aspects of the digital footprint analysis.

**Social Graphs and Social Networks**

The notion of graph arose and is used in discrete mathematics. According to Wikipedia, the social graph reflects the links among individuals on the internet networks which ought to describe the relations of the real world. In reality these notions intertwine and now social networks are most frequently comprehensible as communications on the internet. In this paper, the social network notion will be used in the sense of internet communications (most popular are: Facebook, LinkedIn, Twiter, Youtube, Skype, Flickr, hi5, Klase.lt, One.lt). The initial data presented in a social network embrace not so many aspects: hobbies, activities, social and demographic data (Fig. 2). By adding the whole digital footprint (friends, influences, games, contents under development and much more from the internet and TV), we get a powerful real time data flow bearing plenty of information about an individual and his environment. The nature communications social networks in Fig. 3 are disclosed with humor by one of the authors of the portal Nežinau.lt.

**Social Medium**

This is a phenomenon of the late age which arose after developing easily accessible and simple online published technologies used for social intercourse. Social media use internet technologies that turn the communication into an interactive dialogue. Kaplan and
Haenlein (2010) define social media as a set of created applications based on Web 2.0 ideology and technologies that enable us to exchange the contents developed by users. From the standpoint of business the contents developed by users is stressed more. Some of the authors (Wedergem, 2011; Fuchs, 2010) also notice the wrong side of an active role of user. It seems to them that the users’ contents generated in youtube, MySpace or Facebook leads to new forms of exploitation and work reorganization in information capitalism. In all definitions amalgamation of technology and social interactions is common in creating a new value by common efforts. When estimating the situation the founder of internet Tim Berners-Lee (2010) expressed his concern as to unlimited spreading of social networks and emerging monopolize tendencies. The statistic of social media spread and usage in startling large data volumes are accumulated and their usage is sometimes illegal. Social media systems can accumulate and analyze activities of participants even without their suspicion. The data flow is constantly received from users and this large data arrays are accumulated that include individual interests, groups, behavior and their
location. Discrimination between the content and flow data was clear in the traditional communications processes, e.g., in telecommunications the flow data are: who phoned, where it was phoned, how long it lasted and the content data are topics of conversation. In the case of internet, such discrimination is complicated or maybe impossible at all (Štatilis, 2005). To ensure the respect for private life in e-communications by controlling that communications for the sake of lawfulness, there is a lack of some legal acts following the laws, edition of the Lithuanian criminal code has some drawbacks in sanctifying information retrieval from service providers about the past e-communication events (Štatilis, 2005). Thus, juridical problems are important and not yet solved. This is the reality to be understood and accepted, because, just like the whole democratic world, we have regulations for developing an open and free contents and exchanging it on the internet. Following the vivid statement of Eurocommissar Meglena Kuvena (Behavioural Targeting . . ., 2009): “personal data are the new fuel for the internet and the true currency of digital world”. There are attempts to retain the accumulated information unidentified though to personalized, to inform the subjects of private data about possible actions and to get permission for using their data, however they are not always a success. The efforts to control the process in various countries of the world are overviewed in the next section.


Let us look over the attempts to solve data privacy problems that emerged with the development of the new technologies in several countries.

The research data of internet users show (Story, 2008) that large internet companies now learn about their customers much more than at any earlier time: what they are looking for, what is their taste and choice, because hundreds of visits per month of a typical user are fixed, users do not complain of data fixing as yet since most frequently they are not aware of that: they are not informed or simply they do not read the service providing conditions expressed in a hardly understandable lingo. Defenders of privacy tried to sound the alarm time and again referring to the large amount of data aggregated by companies. According to data of one investigation (Samuelson Low, Technology & Public Policy clinic at the University of California at Berkeley), 85% of grown-ups in California would not allow to watch their behavior on the internet in order that a company give the most relevant advertisements to them. By the data of Berkeley university, the majority of the USA web users rejected “behavioral advertising” (US web uses . . ., 2009). The University of Ottawa Canadian Internet policy and Public Interest Clinic (Top Story, 2008) has required exploring the situation as to the usage of internet users giving them advertisements. Since 1999 a group of internet and data handling companies were concerned how to change the data collection technology, but they were not a success. The USA State office of users’ rights protection, Federal Trade Commission (FTC), has applied the self-regulation method since 1990 to guarantee that a user would get information about a possible usage of his data for target marketing (Protecting Consumer . . ., 2010). This method turned out to be insufficient and now a new law is prepared. It was evident that to
ensure the internet privacy, the so-called “notice&consent” approach not sufficient. After discussions in 2009–2010, FTC has stressed to have found in the present society the followings things:

- universal collection and usage of user’s data;
- growing flow of information and computer power;
- tender limits between personal and anonymous data;
- importance of privacy for many users;
- lack of understanding and limited opportunities for a user to select a safe strategy.

FTC is preparing a new strategy to be announced in 2011. The European Commission (EC) is also concerned about the current interaction problems. Interaction defines the ability of computers and information systems to exchange and use one another’s information (Digital agenda, 2010). EC intends to rewrite the rules of the IT sector with a view to:

1. Increase the users’ trust.
2. To set the “Digital agenda” that regulates interaction and standards.

EC will revise the current Data protection directive. The executive power of EC promised to announce the results already at the beginning of 2011. One of the regulations of the directive is that the personal data “have to be stored no shorter than 6 months and no longer than 2 years since the date of communication”. Up till now data storage terms are very distinct in the EU counties. EC promise to control how the internet companies, especially social web pages such as “Facebook” and advertisement companies use personal data of citizens. Other participants of the digital agenda discussion think that “this directive is the measure, violating the privacy at most in it extend, that has ever been adopted by EU”. The strategy prepared states that “citizens have to be informed what rights they have to access, correct or delete their data”. On the other hand, the right to delete their digital footprints may turn into a possibility for criminal elements to delete their crime footprints ... there was a desire to supplement the “Digital agenda” with the statement that supports internet blockage and filtration, however, for other people these objectives cause a great concern. There are more problems than real solutions in this field so far.

4. Benefit and Possible Threats of Social Media in the Context of Participants in the Process

All those participating in the process of social media cooperation, mentioned in the introduction, have their own challenges, benefit, threats and dangers about which we can discuss only in a concrete context. We have elucidated peculiarities of the data collection and accumulation process and the importance of a digital footprint when analyzing and grouping the social internet participants. The expected benefit, advantages, and possible pitfalls threats action of interdisciplinary fields are presented a summary table “At a glance” (Table 1).

Internet and social media are an unknown area to many governments and legal institutions, because they merely do not know what to start from. Several counties try to
regulate data preservation by their own laws, however the e-networks is global – without bounds and territories, other that the way of life usual for mankind for ages. There is no common legal system in the global world and it can hardly ever be. Let us look at a typical situation: a citizen of country A is looking for information in sever of country B, data were collected in countries C, D, and they will be used, for instance, in country E for users’ behavior advertising. Of which country would the laws be in force for data protection? Traditional regulations of the intellectual property right do not suit in the case of the global network. Possibilities of cybernetic attacks and experiences have exposed dramatic violability of network systems and lack of protection strategies. Dangers in the networked global world are increasing and self-control, based only on notification and a user’s approval of personal data collection is ineffectual. Attempts to regulate data privacy matters were not efficient so far, there are no proper laws, protection technologies lag behind and are unable to ensure the security of systems.
In view of the situation I can see three ways of solution that would cover all the participants of the process:

(a) users,
(b) companies of social media,
(c) specialists creating software for social network maintenance, data collection and analysis.

The first way is education of a deeper user’s understanding about a digital footprint and how it can be used for his disadvantage. To this ends, it is necessary to acquaint users much more with the existence of its digital shadow and with all the possibilities of its utilization, which would help a member of e-society to avoid many troubles and pitfalls of digital media. The second important factor would be honesty of the companies that collected data analyze and use them to attract their customers – users of their services and products. However we can hardly trust the honesty of companies that are mostly oriented only to profit. Very strict laws would be indispensable that are almost impossible in the global network, as we have already found out. On the other hand, even the strict constraints will hardly change anything in essence if powerful profit interests dominate – some ways will be found how to evade any restrictions. It remains the third most important leverage on stopping questionable collection, analysis and usage of personal data. That is a clear understanding of responsibility of the specialists who create information systems for interactive collaboration on the internet and legalization of professional ethics norms in information technologies. IT specialists are one of the most important links in developing social networks. Marketing technologists could surely not accomplish anything without the help of informatics specialists. These have to know clearly what data are collected, where and for what they will be used, what consequences of that data analysis could ne and they have to draw adequate conclusions. Just like medical professionals can say “no” to some procedures, biotechnologists obey the ethical principles in their research, the specialists of engineering informatics have also approached the borderline in the current development of their sciences when, in certain cases, they have to say: “No, I will not do that”, based on their professional ethics code. Codes of professional societies are no news; they were created and followed as far back as the past age. Now it is time to recall and renew them by adapting to the up-to-date technologies, this activity is preceded rather intensively: the scientific ethic is promoted in the USA and Europe, and the Global Ethics Observatory sponsored by UNESCO is operating.

5. Codes of Professional Ethics

We have elucidated that there are no universal methods that ensure data privacy of the internet browser or other social media participant and that allow only enjoying the advantages of the new technology. One of the few possibilities in that all the participants engaged in the computer network process, those collecting data, analyzing them or constructing programs necessary for marketing would observe their Professional Ethics Code. Scientific and professional ethics became especially urgent due to the wild development of the new IT attacks of computer virus, hacking into computer systems would
be impossible, if everybody obeyed the ethical principles. Computer viruses are created under the semblance of security systems that attack and threaten users, as shown in Fig. 4. Such programs can be developed only by professionals.

There would be no such cases if IT professionals followed the professional ethics code. Let us consider the heritage available in the field of ethical codes. The oldest code that will soon celebrate its centenary – Code of Principles of Professional Conduct of the American Institute of Electrical Engineers, is dated 1912 (Code of Principles...). Now it is accessible via the digital library of the Professional ethics Study Center (Online Ethics Center...). The general principles, consisting of 22 of them, remain the same in the lapse of time, starting from: “1. in all the activities an engineer should follow the topical principles of honor” and finishing by: “22. An engineer responsible for work must prevent a change of his engineering decision for persons without a technical education”. The library mentioned has aggregated lots of professional ethics codes. Nowadays the Online Ethics Center together with National Engineering Academy is operating in the Illinois Technological Institute. They provide easily accessible information on engineering ethic topics issues meant for studies and discussions – both for practitioners and educators, for students and all individuals taking interest in the professional ethics of researchers. One can find many scenarios that concern internet privacy issues relevant to us as well and suitable for teaching. Such a course would surely be of use at Lithuanian universities that prepare IT specialists. Even mathematicians in America have their ethical principles: Ethical Guidelines of the American Mathematical Society, 2005, though they seem to have little opportunities to violate the ethics in their research. The Code of Ethics of Association of Engineering Technicians and Technologists of Newfoundland and Labrador, Canada, 2009, are quite modern. It contains, e.g., such principles: “To act jointly with respect to customers and employers” interests, to observe confidentiality and avoid the conflict of interests, and in case it happens, to disclose the circumstances immediately both to a customers and an employer”... “To work honestly, to communicate with customers, colleagues and others with respect and confidence, to appreciate the work of
others, to accept and honestly make right notes.” Such principles would be also applied in the IT professional ethics code in Lithuania. It is of importance to pledge ourselves not to conceal possible consequences of data collection and their processing, which are often understandable only for specialists and usually inconceivable for ordinary media participants.

Researches in the scientific ethic field were pursued in Europe as well. In 2000, the European Science Foundation initiated a discussion in the society of researchers on the application of the highest standards in scientific research practice. As a result, a survey has been published (European Science . . . , 2008) that considers the situation in detail of 18 countries (out of 32 countries) it disclosed the examples of good practice in scientific research as well as misunderstandings still present in different European countries. The survey offers recommendations to scientific societies, to agencies that financially support science and to research institutions, and presents examples of good and bad practice, that could be applied avoiding some dangers, arising in interdisciplinary researches of social media. Another significant source for research and practice of ethical problems is a database system developed and supported by UNESCO. The Global Ethics Observatory (GEObs, 2005). It is a repository of rules, ethical codes and teaching programs of all the world of tree access, searched by ethical institutions and experts, operating since 2005. It consists of 5 independent databases: of ethical experts, institutions functioning in the ethical field, programs of teaching ethics, ethical codes of law and, finally, codes of ethics – it contains 151 codes of ethics. After analyzing many sources mentioned above, we can pose questions to ourselves: whether the Lithuanian IT specialists can be responsible for their works, can behave following the professional ethics code, and can work not only for career or money.

6. Conclusion

With a violent change in IT and their tools both the rules of theoretical research and methods and IT tools and applications have to change as well. A massive spread of communications, created by users they make an interesting space for original research in a computer networked society and many critical remarks appear thereby. This paper has reviewed the positive and negative features of the new phenomenon – social media.

Positive Aspects

When analyzing every stromily changing situation during the past two decades, the positive changes observed are as follows: former users become creators of digital contents; their inclusion is increasing. After analyzing users’ needs, the companies are working more efficiently: personalize the supply, develop only what the users need, innovative products and services, fit to the internet visitors’ needs, come into being. A large amount of data accumulated is the basis of research after the analysis of which behavioral targeting is undertaken, and only the needed and interesting services and products are proposed to customers.
Negative Aspects

In reality the advantages mentioned above rather frequently are derogated by means of malicious usage of data and information. There is additional risk in this field of research due to great penetration of internet in the society and a large amount of users involved. Temptation of illegal data usage and easiness arises due to that individuals and companies often are not aware what is fixed and how their data are used and analyzed, where they are stored and transferred. Most frequently there is no possibility of controlling or eliminating our own data that have got into the digital space. The measure of legal regulation lag behind, the law and order officers often do not conceive that the e-environment needs essentially other measure that in the traditional environment. Securities technologies cannot keep up with the capabilities of the new media, even the bank of Information System considered safe so far are also breakable. It suffices to recall the recent hackers’ attacks of PayPal systems in connection with WikiLeaks scandals.

Responsibility of Researchers and Tasks for the Future

It is necessary to develop security technologies in order to stop malicious usage of data and information. There are no methods as yet how to surely evaluate the social network security. Just like the quality of digital repositories is estimated (Kligienė, 2009), it is necessary to develop the methods and criteria for estimating a more complicated phenomenon, i.e. to estimate the quality of social networks – to find aspects that predetermine the quality, to find proper criteria taking into consideration a great many of participants involved in interdisciplinary research and practice. The measure for observing social media are just begun to develop (Chinwag ,..., 2008; Measurementcamp, 2011). It is needed to develop methods suitable for data research other than traditional. The part of researchers when creating the rules and applying the codes of professional ethics is of almost importance. We can diminish threats and dangers as yet only by enlightening users, educating specialists in the spirit of professional ethics already at the university lecture-rooms, by stressing the importance of professional ethics in the stages of social media creation and keeping to them. Thus, social media pitfalls that threaten users because of possible usage of digital footprints could not be dangerous in the context of professional ethics.

References


S.N. Kligienė


S.N. Kligiënė (1944.03.03 – 2011.06.18) has been with the Institute of Mathematics and Informatics since 1967. She graduated from Vilnius University in 1967 in mathematics, and she obtained her doctorate there in 1973. She was an associate professor at Vilnius Gediminas Technical University and lectures in mathematical statistics and time series analysis. Her scientific interests were statistical analysis of time series and the detection and estimation of changes in model parameters. She is the author of more than 70 scientific articles.

Skaitmeniniai pėdsakai internete profesinės etikos kontekste

Stanislava Nerutė KLIGIENE