The Privacy Paradox in the Context of Online Health Data Disclosure by Users

Chrysanthi Kosyfaki[1], Nelina P. Angelova[1],

Aggeliki Tsohou[1], Emmanouil Magkos[1]

[1] Ionian University, Department of Informatics
Tsirigoti Square 7, 49100, Corfu, Greece

{p13kosy, p13angk, atsohou, emagos}@ionio.gr

Abstract. The privacy paradox phenomenon corresponds to the inconsistency between the privacy concerns and the actual behavior of online users. The existence of the phenomenon has been studied in many fields, such as for social networks and especially for a variety of forums and online communities. This paper questions its existence in the context of sensitive data, and more specific in the health data area via a survey which took place in Greece. Given that health-related information is perceived as sensitive personal data it is often excluded from discussions and sharing. This research aims to unravel the paradox's existence regarding the disclosure of health data information when individuals visit related online forums and communities.

Keywords: User privacy, privacy paradox, health data, data disclosure

1   Introduction

The past decade has been characterized by the explosion of the web and online social networks. The users' tendency to disclose personal information, together with their expressed concerns about their privacy's protection, raised the phenomenon of the privacy paradox. People commonly give away their belongings only when an expected gain coming from this act, and this gets bigger than the profit of keeping them. When this does not apply, we are dealing with the paradox phenomenon.

With the technological evolution, nowadays, there are plenty of online sites and communities that refer and discuss about any kind of information. Therefore, there are many sites in which users can search information, share problems and thoughts, and ask others for help. The pitfall is that in order to participate in this kind of conversations, one ought to talk openly about his/her issue, and disclose personal sensitive data of him/her. In Greece, it is common for people to visit sites and online communities where health issues are being discussed. Some examples of such sites include Mammyland[1] and Iatronet[2]. The members of these communities communicate

---

[1] http://mammyland.com/

through comments, ask questions about their problems and they give specific information about their health status and their personal details, thus disclosing much about their health condition and their identity.

Our Contribution. This paper aims to reveal if this kind of disclosure of medical data seen in forums and online communities, reflects a privacy paradox existence. In order to explore that, we developed and conducted a web survey in three stages [11, 12, 13] with participants from Greece. To the best of our knowledge, this is the first time that this issue is being studied in the context of health data. The results of our study question the existence of the privacy paradox phenomenon and show an emotionally driven behavior from the side of the users.

The paper is structured as follows: In Section 2, privacy and health data definitions are being given, while in Section 3, the privacy paradox phenomenon is being discussed. In Section 4, the results of our study are presented, while in Sections 5 and 6, the results are discussed and conclusions are drawn.

2   Privacy and health data

Privacy, as Alan Westin contends, *"provides individuals and groups in society with a preservation of autonomy, a release from role-playing, a time for self-evaluation and for protected communication"* [1]. Private data is defined in the 1998 Act[3] as "*data which relate to a living individual who can be identified – (a) from those data, or (b) from those data and other information which is in the possession of, or is likely to come into the possession of, the data controller, and includes any expression of opinion about the individual and any indication of the intentions of the data controller or any other person in respect of the individual"*. Sensitive data are private data that are capable of revealing information about ethnic or racial origin, political opinion, religious or other similar beliefs, memberships, physical or mental health details, personal life and/or criminal or civil offences [7, 8]. Health data on the other hand, is defined as any information used to provide, manage, pay and/or report on the services used across the entire healthcare system [9].

3   Privacy paradox and health data disclosure

The *privacy paradox* suggests that while Internet users are concerned about privacy, their behaviors do not mirror those concerns [14]. Many researchers have studied the privacy paradox phenomenon and a single emerging development to explain this has yet to emerge. The first attempt to understand this phenomenon was in 2006 by Barnes et al. [3]. In their study, they mainly explain the teenagers' behavior in online social networks and especially the information that they choose to share with their online friends. Furthermore there are several studies that explore the existence of the phenomenon in other fields of social sciences [16], like Norberg et al. [9], who eventually established the definition of privacy paradox in 2007 [6], [9]. Literature provides insights that the existence of the phenomenon relates to the users emotions. In 2014, Kern et al. [5] tried to prove that the emotional state of users is responsible for the existence of the phenomenon. Moreover, another research by Fleisch et al. [15] adopt the same approach. Our findings are in line with this approach and provide evidence that emotional factors are significant for the existence of the phenomenon in the context of health data.

While researchers study the privacy paradox phenomenon, it is known and generally accepted that individuals increasingly use information blogs, social networks and personalized applications for various reasons, such as curiosity and socialization. In

---

[2] http://iatronet.gr/
[3] https://www.privacy.org.nz/the-privacy-act-and-codes/the-privacy-act/

the case of health data, people have organized online communities and forums to exchange opinions and knowledge, to discuss their health issues and support each other. The owners of these forums and web pages seek to collect and manipulate individuals' health data in exchange of their services. While many researchers are trying to tackle with this situation and question the existence of the phenomenon in the field, two questions are naturally raised [2],[4]:

- *Are people aware of the importance of their health- related data?* If one cannot understand the importance of this kind of own data, then he/she will not be able to efficiently value and, consequently, protect it. Do people who disclose their health-data acknowledge their importance? Do they know the processing and handling that their data may go through by those who own the copyrights of the data posted and shared through their websites? If they don't, then we cannot talk about a privacy paradox phenomenon. If they do, a second question needs to be answered.

- *How individuals assess that value of their health data?* If the value they assign reflects their online behavior, then we cannot talk about a privacy paradox either. The only case where a privacy paradox's existence exists, is if people's disclosing behavior does not align with the value that themselves assign to their data.

4   An empirical study

In an attempt to answer the questions made above, understand what forces people to participate in health-data disclosing discussions and compare the gain of this participation to people's value about their data in order to come to a conclusion about the phenomenon existence, we conducted a survey in Greece, which included the invitation of participants who are users of social networks such as Facebook as well as members of online communities. Our sample included 163 participants. Most of them (about 74,6%) belong to the age-group of 18-40.

When participants were asked if they know what kind of their data is categorized as sensitive, 76% claimed that they do. But, when they were asked to shortly name the categories they think that fall into the space of sensitive personal information, there were few answers in which health data were mentioned (see Table 1). These answers indicate that almost 25% of the participants didn't have any idea about what is sensitive data . Our results coincide with other studies, that also show unawareness and ignorance from the side of the users [4].

Table 1. Answers to the question of which they believe that personal data are

Percentage	Answers
63.00%	Properties and belongings
14.00%	Credit card PIN Codes
10.00%	Medical Data
9.00%	Religion beliefs
4.00%	Criminal Records

When asked how important the privacy of their health data information is to them,

although many claimed that this kind of protection is indeed of primary importance, 1/10 did not seem to agree. In order to understand what makes people participate in online forums and look for information about their health, we asked how many of our participants had ever visited a web page or forums about health information. Among the participants 65% said that they have visited such sites at least once in their lifetime in order to search for information about health data issues. When asked about the reasons that led them to this kind of online behavior, their answers varied but they all had something in common: they included some kind of emotional factor (see Table 2).

Table 2. Reasons for visiting health forums or online communities

Percentage	Answers
57.00%	Quick Information
20.00%	Anxiety/Fear
10.00%	Curiosity
9.00%	Conversations with Homeopathic
4.00%	Valid Diagnosis

Those who have visited these web pages were additionally asked if they had ever participated in a discussion with other members of the online community or forum. The majority of the subjects (90.1%) declined that they have ever shared and posted information about their health issues. Furthermore, the participants were asked about the way that they posted or visited this kind of sites (i.e., using username, using a nickname or anonymously). The same question was made hypothetically to the participants that have never visited and posted in such communities. They were particularly being asked, which of the three options they would choose, if they were to comment and participate in online conversation: anonymously, using a nickname or by giving their true personal information (see Figure 1). They were asked to justify the reasons behind their choice (see Table 3).
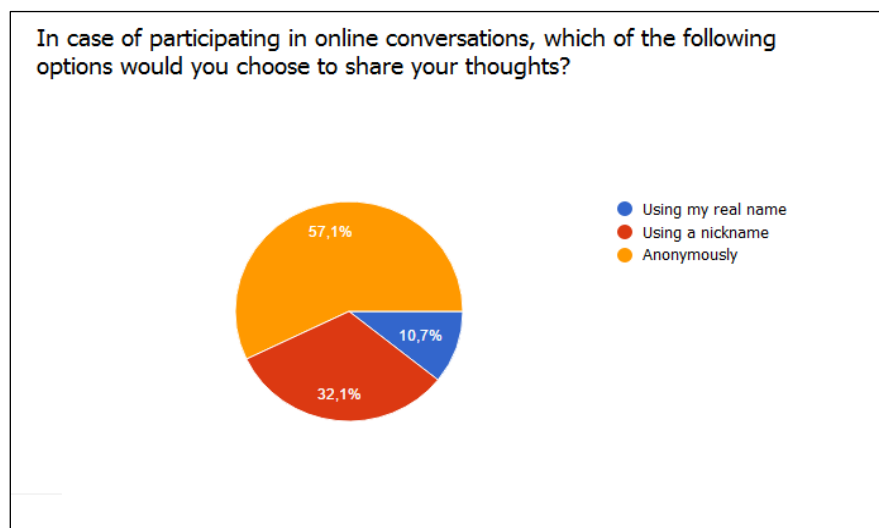

Fig.1 Participants' behaviors with regards to personal identification (source: our survey)

Table 3. Elaborated replies on choice selection (real name, nickname, anonymity) with regards to personal identification

Answers
I don't think that it is necessary to disclose my true personal details
I would prefer to participate anonymously in order to avoid an association of my comments with my identity
Conversations where real names are being used tend to be more serious
I don't regularly use my real personal information, I am trying to avoid it when possible
I'm ashamed
My identity does not change the diagnosis
I have nothing to hide
For my identity's protection

From these answers, it can be derived that the majority (57.1%) of the participants have concerns about the protection of their identity and privacy, and they would prefer anonymity in order to feel more protected.

70% of the participants that have posted, answered that they used a nickname, 20% used their real name and 10% were anonymous. But when we asked them if it was necessary to disclose personal details their answers were positive. The majority of people said that they gave their email to create an account to those pages and without this, the login was impossible. The participants were asked to describe the reasons that they preferred nicknames and they believed that it is a way to protect their privacy (Table 4). However, when asked if they feel safer by using a nickname, almost 85,7 % said they do not (see Figure 2). Moreover, in the answer about which of the data that they have given is real, almost everyone said that they have given their true personal information.
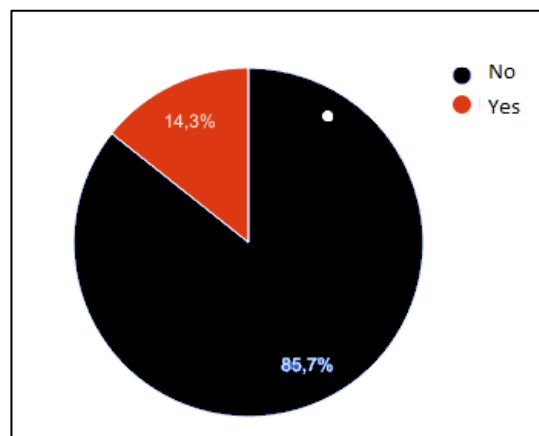

Fig.2 Do you feel safer by using a nickname? (source: our survey)

From the survey we understand that users are confused about which is the most appropriate way to post their thoughts (about 42.9% believe that it is safer for someone to use their real name while almost the same percentage has the opposite opinion).

Table 4. Reasons why users prefer nicknames

Percentage	Answers
70.00%	Privacy
18.00%	It is not necessary for someone to give his/her real name
9.00%	Possible recognition
3.00%	Unwilling to giving my real name

Those who claimed that they have shared their problems asking for help, either from other members of the online community or from specialists (doctors), were asked what do they think that they gained in return from this kind of disclosure and sharing. Their answers were once again driven by their emotions (see Table 5), in line with what other research's results indicate [5].

The participants were also asked to compare their perception about the privacy loss caused by their participation to the online sites with the gain that they think that this offered them in a return. With this question we actually called the participants to determine the existence or non-existence of the privacy paradox phenomenon in the field. If their gain makes the disclosure worthy, then we cannot talk about a privacy paradox, as people seem to disclose information about their health conditions for something fair in return. From the responses 65% of those who answered that they have commented and exchanged opinions in this kind of social forums feel that it was definitely worth the price.

Table 5. Replies regarding the gain that individual perceive to take from their visit

Percentage	Answers
52.00%	Relief
25.00%	Useful opinion exchanging
12.00%	Courage
7.00%	Security
4.00%	Additional anxiety

5 Discussion

Many people who participated at our survey, claim that they value and care about their sensitive data, like the data exposing their health status, conditions and issues. Nonetheless, they give away such information in forums, communities, professionals' web pages and applications. Some of the questions that our study tried to address related to what leads users to this behavior and the participation in those activities. What do people gain in return so that they assume that it reflects the "label of importance" that they put upon their health data? Answering such questions, brings us one step closer to understanding the privacy's paradox phenomenon and question its existence. Here, we have two possible situations that would prove the nonexistence of a privacy paradox in the field: either the gain seems very high from the point of view of the users (i.e., high as the value the assign to the data), or their health data is not of much importance to them. As the beliefs of the users are a key to answering this question regarding the existence of the privacy-paradox phenomenon, our empirical study has shown the following scenarios:

- **Scenario 1:** We are dealing with many online users that do not consider their

health data as sensitive, due to lack of awareness, ignorance or misinformation. Given that these users do not assign high value to their data they are not making great efforts to efficiently protect it. This leads us to the conclusion that we cannot talk about a privacy paradox phenomenon, as users don't understand the importance and value of their data and they are not putting them in danger knowingly.

- **Scenario 2:** For those who value their data and know the data's importance, we can see that their online behavior is emotionally driven. The most frequent visitors to forums were women (52%) who were worried about their health or their relatives' health. Therefore, we cannot talk about a privacy paradox, as emotional overload can affect the users' judgment and rational decision making and make them more prone to disclosure.

- **Scenario 3:** If a case does not fall in the above two categories, then we still cannot accept the phenomenon's existence. This is because in the last question made in the survey, the answers of the participants showed clearly that the benefit of sharing personal information is worth the disclosure of health information.

From the findings of our survey, thus, we can summarize and argue that:

The privacy paradox does not exist in the field of
health data disclosure made by online users in communities and forums.

The only case in which we could assume that the privacy paradox seems to exist, is when users said that they had participated in conversations, after they had given away personal details during their registration. But we should take into consideration that they were emotionally driven during the time of the disclosure, something that takes us back to the Scenario 2.

As we mentioned above we are not the first who study the phenomenon from the point of view of users' emotions and emotional state [5],[15]. However, to the best of our knowledge, it is the first time that a study addresses the existence or non-existence of the phenomenon in the health field and contributes to the general research upon users' behavior and the existence of the privacy paradox. We should note that our findings come from a relatively small sample of participants in Greece and they be considered as a starting point giving spark to other similar studies in different countries.

6   Conclusions

In this research we questioned and studied the existence of the privacy paradox when users disclose their health-related data in online communities. To the best of our knowledge, this is the first analysis of the privacy paradox in the health field. The outcome of our research is that the privacy paradox phenomenon does not exist in this context. First, in many cases the users do not seem to understand the value of this type of data and thus they disclose them due to lack of awareness. Second, during the time of the disclosure, they seem to be emotionally driven. When someone is emotionally charged, it is difficult for him/her to make the right decision and behave rationally. As a future work, it would be interesting to analyze this phenomenon globally, to understand whether the privacy paradox phenomenon in this area exists or not, in countries besides Greece.

References

[1] Westin F. Alan: Privacy and Freedom (1968)

[2] Acquisti et al: What is privacy worth?, The Journal of Legal Studies (2013)

[3] Barnes,S.B.: A privacy paradox: Social networking in the United States (2006)

[4] Brandimarte, L., Acquisti, A., and Loewenstein, G.: Misplaced confidences privacy and the control paradox, Social Psychological and Personality Science (2013)

[5] Kehr, F., Wentzel, D., and Kowatsch, T.: Privacy paradox revised: Pre-existing attitudes, psychological ownership, and actual disclosure, Proceedings of the Thirty Fifth International Conference on Information Systems, Auckland, New Zealand (2014)

[6] Lee, H., Park, H. and Kim, J.: Why do people share their context information on Social Network Services? A qualitative study and an experimental study on users' behavior of balancing perceived benefit and risk, International Journal of Human Computer Studies (2013)

[7] MIT: Sensitive Data : Your Money AND Your Life (2008)

[8] Hellenic Data Protection Authority (HDPA), http://www.dpa.gr

[9] Norberg, P.A., Horne, D.R., and Horne, D.A.: The privacy paradox: Personal information disclosure intentions versus behaviors, Journal of Consumer Affairs (2007)

[10] Tzourakis M.: The health care industry and Data Quality (1996)

[11] Google Forms, https://goo.gl/forms/ZXuMRHGeZJA7UJA92

[12] Google Forms, https://goo.gl/forms/upq6BlwSh2SG9VO03

[13] Google Forms, https://goo.gl/forms/3lxKzMI6oiGCraVY2

[14] Web Privacy Wiki, http://web-privacy.wikia.com/wiki/Privacy_Paradox

[15] Kehr, F., Kowatsch, T., Wentzel, D., and Fleisch, E.: Blissfully ignorant: the effects of general privacy concerns, general institutional trust, and affect in the privacy calculus, Information Systems Journal (2015)

[16] Holland H. B.: Privacy Paradox 2.0 (2010)