# Protection of Personal Data in Information Systems

**Bojken Shehu [1], Shqiponja Ahmetaj [2], Marin Aranitasi [3] and Aleksander Xhuvani[4]**

**[1] Computer Engineering Department, Faculty of Information Technology, Polytechnic University of Tirana Tirana,Albania**

**[2] Computer Department, Faculty of Informatics, Vienna University of Technology Vienna,Austria**

**[3] Center for R&D, Faculty of Information Technology, Polytechnic University of Tirana Tirana,Albania**

**[4] Computer Engineering Department, Faculty of Information Technology, Polytechnic University of Tirana Tirana,Albania**

**Abstract**

The rapid development of computing technology has led to the emergence of the greater capacity to store personal data. The huge amount of information that governments and businesses collect from individuals have become a cause of concern. Personal data collection encroaches on the individual's right, particularly as it invades privacy or the right to control information about ourselves; there is a disclosure of private personal facts; and, the information can be used in such a way that a person is cast in a bad light or in a case of identity theft. The method of personal data collection, its use and security, hence, necessitate citizen protection. Through the years, laws that aim to protect personal privacy have enacted but they appear to be insufficient. This paper examines the concept of depersonalization as an alternative method for the representation and protection of personal information. It is based on the argument that the legal protection available is not sufficient to address privacy concerns.

***Keywords:*** *Data Security, Depersonalization, Legal Protection, Personal Data*.

## 1. Introduction

The rapid development of computing technology has led to the emergence of the greater capacity to store personal data. These data are important and collected because of research. Governments collect them for the evaluation of its programs or for use databases such as in law enforcement and social security. Businesses are voracious consumers of these data because they are crucial in the success of several operations because they record habits, preferences, among other patterns of individual activities that help them develop market and sell products and services. Computers and the Internet paved the way for more efficient and fast methods to gather, store and organize personal information. Since the 1970s the number of computer data banks or databases became staggering. There are the databases from Social Security Administration, Law and Order authorities (like the Federal Bureau of Investigation in the USA), Medical Information Bureau, state criminal justice systems, municipal data systems, credit card companies, telephone companies and recently Google, Facebook and so much more.

This paper will explore depersonalization as an effective method of personal data collection, where privacy is still protected. The description and analysis of depersonalization reveal a sound framework that can achieve a level of privacy protection in a manner that does not hinder the need of governments and businesses for statistical research. These data are important and collected for statistical purposes. Governments collect them for the evaluation of their programs or use them as databases in different areas. Businesses are voracious consumers of these data which are crucial in the success of several operations. They record habits, preferences and other patterns of individual activities that help them develop, market and sell products and services. Computers and the Internet paved the way for more efficient and fast methods to gather, store and organize personal information.

## 2. Legal Protection of Personal Data

The foundation of legal protection against indiscriminate collection of private data is a Congressional report that outlined four tenets of fair information practices, namely:

- Notice or the disclosure of the details of data gathering practices, policies and results to data subjects.
- Choice or the ability of the data subjects to exercise choices about how their personal data can be used.
- Access or the level of access provided to individuals on the gathered data about them.
- Security or the responsibility of data gatherers to provide adequate protection for the information collected [1] (Bidgoli 2004, page.98).

Based from these principles, a number of laws were enacted covering individual privacy across different sectors. For instance, the Gramm-Leach-Bliley Act protects personal banking information; the Fair Credit Reporting Act provides the framework for handling personal credit data. There are also laws that cover the collection and use of medical and health data, government records, children's privacy, and so forth. Laws are also enacted in other countries such as the European Union Data Protection Directive, the OECD privacy guidelines adopted by countries such as Mexico, Australia, Japan and Czech Republic [3] (Conrad, Misenar and Feldman 2012, page.401). But these laws and even some ethical guidelines [5] (Kluge 2000), no matter how specific and comprehensive, still fail to address privacy issues. Neubauer and Kolb [7] (2009) , for example, noted that approaches and methods for protecting privacy often do not comply with legal requirements or basic security requirements without suffering any penalty, (7). Szeto and Miri [10] (2007) revealed the same findings when they studied the Canadian experience. According to Hildebrandt and Gutwirth [4] (2008), this is because most statutes builds on traditional ways of thinking data, personal data and their abuse, without understanding or recognizing the new type of knowledge that result from modern data processing (p.321). It was further argued that even when recent or updated laws were effective regarding personal data, they are still not equipped to deal with correlated data, which is persistent today since "(1) group profiles are often inferred from anonymous personal data to which data protection regulation do not apply and (2) group profiles do not necessarily apply to identifiable persons but may, even so, affect the autonomy, privacy, security and equality" of individuals (page.321).

## 3. Depersonalization.

Depersonalization is a concept in personal data collection that builds on the principle that researchers do not necessarily need the personal identities of data subjects in order to be effective or to achieve objectives because what is only required for legitimate research is statistical access. Ideally, depersonalization renders a data subject completely anonymous. However, this is impossible to achieve in most applications that is why a modified definition was put forward, which states depersonalization as "the modification of personal data so that the information concerning personal or material circumstances can no longer or only with a disproportionate amount of time, expense and labour, be attributed to an identified or identifiable individual" [2] (Fischer-Hubner 2001, page.112). The definition was contained in the groundbreaking Federal German Data Protection Act, which already became synonymous to practical depersonalization. To demonstrate this in real-world application, this paper cites a Lightweight Data Security System developed by Rawassizadeh [9], which provided a working framework that can provide insights how depersonalization actually works (Fig.1). Based on the architecture, the user inputs data into a system that include several stages of pseudonymization before personal information is published or made available to third parties.

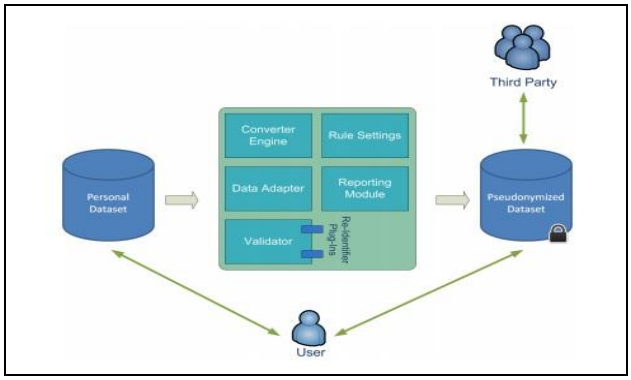



Fig 6. Lightweight Data Security System Conceptual Framework (Rawassizadeh at al ,page. 3)

Neubauer and Heurix [6] (2011) further provided specific applications when they proposed a depersonalization system or pseudonymization of medical data to be used in health care institutions. The application is called Pseudonymization of Information for Privacy in e-Health or PIPE, which aims to provide a "traceable anonymity" (page.194). It works using a combination of symmetric and asymmetric cryptographic keys in order to achieve a logical multi-tier hull model composed of three layers

(Fig. 2). The model is applicable in several health care scenarios as shown in Fig. 3.

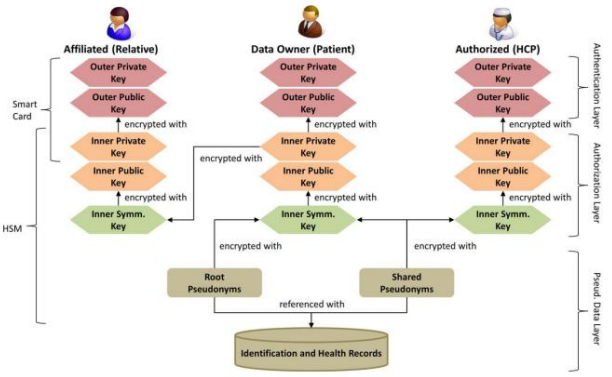
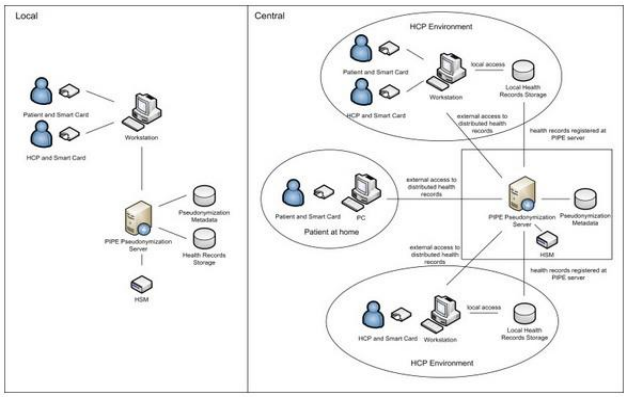



Fig 2. The PIPE Model (Neubauer & Heurix, 195).

Based from the sample frameworks provided, it is clear how the depersonalization system works and how data is stored, maintained and protected. Its adoption will entail the installation of additional application but it will effectively address ethical and legal questions on personal data gathering practices.




. Fig 3. Scenarios for PIPE Application (Neubauer & Heurix, page.195).

## 4. Challenges and Future Trend

Depersonalization, certainly, is not perfect. The method and its resulting applications still entail risks since the anonymity it could provide is not absolute. This was earlier cited with the modification in the definition of the concept. Actual depersonalization of data, wrote Wagner [11], is weakened by the practice of linking different records to increase their information content and link the outcome to the identity of individuals (page.3). Neumann [8] also argued that the method of depersonalization may lead to temptation to commit misdeeds, diminish human initiative and hinder the principle of accountability, and eroding our sense of ethical behaviour in the process.

There are risks involved such as the methods available to "re-identify" anonymous individuals. But the process, as has been pointed out by the German law cited in this paper, becomes time and resource consuming that undertaking them becomes impractical. There are also mechanisms and applications that can prevent the risk of re-identification. This is particularly important in the area of future work in this field. Like any form of technology, it rapidly evolves, updating capabilities so that there is a potential of more sophisticated and effective models.

## 5. Conclusions

All in all, depersonalization is an effective and viable solution to personal privacy concerns amidst the tremendous power available to governments and businesses to gather personal data. It is a technical solution to a technical problem. The inefficacies of enacted laws to protect personal privacy serve to highlight this point. Fundamentally, adopting it makes sense because it addresses the problem from the very beginning: the identities of data targets are masked and the statistical information is accessible. It solves privacy issues and satisfies the need for data so that governments and businesses are able to provide products and services that are better and more attuned to our needs.

## References

[1] Bidgoli and Hossein, "The Internet Encyclopedia". Hoboken, NJ: John Wiley and Sons, 2004.
[2] Simone Fischer-Hubner and G. Fischer-Hubner, "It-Security and Privacy: Design and Use of Privacy-Enhancing Security Mechanisms " Issue 1958. Berlin: Springer, 2001. Print..
[3] Conrad, Eric. Misenar, Seth. and Feldman, Joshua."CISSP Study Guide". Waltham, MA: Newnes, 2012. Print..
[4] Hildebrant, Mirreile and Gutwirth, Serge. "Profiling the European Citizen: Cross-disciplinary Perspectives". Berlin: Springer, 2008. Print.
[5] Kluge, Eike. "Professional codes for electronic HC record protection: ethical, legal, economic and structural issues." *International Journal of Medical Informatics*, 2.1 (2000): 85-96. Print..
[6] Neubauer, Thomas and Heurix, Johannes. "A methodology for the pseudonymization of medical data". *International Journal of Medical Informatics*, 80(2011), 190-204. Print..
[7] Neubauer, Thomas. and Kolb, Mathias. "Technologies for the Pseudonymization of Medical Data: A Legal Evaluation". 2009 *Fourth International Conference on Systems*, 2009. Print.
[8] Neumann, Peter. "Consideration on risks in using computing technology." *ACM SIGSAC* Review, 6.2 (1988), 2-4. Print.
[9] Rawassizadeh, R., Heurix, J. Khosravipour, S. and Tjoa, A Min. "LiDSec: A Lightweight Pseudonymization Approach for Textual Personal Information." Vienna: Vienna University of Technology.

[10]Szeto, M. and Miri, A. "Analysis of the Use of Privacy-Enhancing Technologies to Achieve PIPEDA Compliance in a B2C e-Business Model." *Eighth World Congress on the Management of eBusiness*, 2007. Print.

[11]Wagner, Gert. "Autonomous Organization of the (International) Scientific Community Would Simplify Data Protection in the Social Sciences and Encourage Reanalysis." *ECONSTOR*: Discussion Paper N. 249, 2001. Print.

**Bojken Shehu**. He is a pedagogue in Polytechnic University of Tirana, Faculty of Information Technology, in Computer Engineering Department. In 2007 he has finished the Bachelor Thesis in Saint Petersburg State Polytechnical University, Russia. In 2010 he has finished the Master Thesis in Bauman Moscow State Technical University, Russia, with excellent results and now he is a PhD student in Polytechnic University of Tirana, Albania.

**Shqiponja Ahmetaj**. She is a Master of Science student in Vienna University of Technology, Faculty of Informatics. In 2011 she has finished the Bachelor Thesis in Saint Petersburg State Polytechnical University, Russia, with excellent results.

**Marin Aranitasi**. He is a pedagogue in Polytechnic University of Tirana. He has finished the Bachelor and Master Thesis in Polytechnic University of Tirana, Albania and now he is a PhD student in Polytechnic University of Tirana, Albania.

**Aleksander Xhuvani.** He is a pedagogue in Polytechnic University of Tirana, Faculty of Information Technology, in Computer Engineering Department. He has finished the PhD study at Bordeaux in France. At 2004 he is graduated as Prof.Dr.