Preparing Students for the Era of the General Data Protection Regulation (GDPR)

Maja Gligora Marković [1], Sandra Debeljak [2], Nikola Kadoić [3]

[1] *University of Rijeka, Faculty of Medicine, Department of Medical Informatics, Braće Branchetta 20/1, Rijeka, Croatia*
[2] *Polytechnic of Rijeka, Business Department, Trpimirova 2/V, Rijeka, Croatia*
[3] *University of Zagreb, Faculty of organization and informatics Varazdin, Pavlinska 2, Varaždin, Croatia*

Abstract – **One of the main goals of the General Data Protection Regulation (GDPR) is to protect the personal data of individuals. Each organization (company, association, school, institution, university, etc.) has an obligation to protect all of the individual data that it obtains. Those data can belong to employees, members, students, clients, etc. The research in this paper is related to the higher education students in Croatia.**

This study is being conducted in three parts. The first part was conducted in April of 2017 (N=159) and the second in April/May of 2018 (N=141), in a period before the GDPR became valid (May 25th, 2018). In this paper, we are analysing the results of the second part of the study. Additionally, we are discussing risks that might appear if students do not know the GDPR. Risk matrix results are used to represent a basis which higher education administrations can utilize to make corrective decisions. The main conclusion of the research is that there are still issues with understanding the basic concepts of personal data and the GDPR, which may cause some problems during studying process. The main recommendation for HEIs or students organizations (such as student councils) is to organize lectures and workshops related to the GDPR.

Keywords – **Personal data, personal data protection, GDPR, virtual environment, students, higher education.**

1. Introduction

With new technologies creating virtual realities, the ability of individuals to connect and exchange information has become almost limitless. The users input their personal data into many websites, which are stored in various databases around the world. However, this creates the potential for the malicious use of data, which requires appropriate criminal treatment.

The European Union (EU) regulation 2016/679 (GDPR), adopted by the European Parliament and EU Council on April 27th, 2016, deals with protection of individuals in terms of personal data processing. Each organization (company, association, school, institution, university, etc.) has an obligation to protect all of the individual data that it obtains. Those data can belong to employees, members, students, clients, etc.

Regarding the GDPR, higher education institutions (HEIs) must make some changes and adopt new processes. In this paper, we are mainly focused on the students' experience of the GDPR; whether they know the basic elements of the regulation, what risks might appear, and how to decrease those risks.

The study is being conducted in three parts:

- The first part was conducted in April of 2017 [1].
- The second part was conducted in April/May of 2018.
- The third part will be conducted in April of 2019.

The first part of the study (N=159) was conducted among students of one Croatian HEI, and the results show that the majority do know how to define the term *personal data* and recognize its forms. Additionally, students of Information and communication technology (ICT) studies give higher

attention to personal data than students not involved in ICT.

The second part of the study (N=141) utilized students from the same HEI. This phase contained nearly identical questions as the first section but included queries related to the GDPR, its application, and the mechanisms of personal data protection it provides. The results are presented in detail below. The risk analysis, presented after the questionnaire, provides recommendations to HEI administrations on how to improve the current state of student knowledge pertaining to the GDPR.

Finally, the third part of the study will be conducted in 2019 among the same HEI students previously used and other students from HEIs in Croatia. This phase will contain similar questions, with the intent of comparing knowledge of personal data and mechanisms for its protection from the very beginning of the GDPR to one year after its implementation.

This paper is organized as follows. In Section 2, the basic definition of the GDPR, with a review of the regulation in higher education, is presented. In Section 3, the research methodology of the second phase of the study is detailed. In Section 4, the results of the questionnaire are discussed. In Section 5, risks and student problems related to the GDPR are theoretically examined by using the risk matrix method. Risk matrix results are used to represent a basis which higher education administrations can utilize to make corrective decisions. Finally, we will present our conclusions.

2. General Data Protection Regulation (GDPR) in Higher Education

Privacy and personal data protection are two interrelated terms that are often used in terms of GDPR. "Privacy generally refers to the protection of an individual's "personal space", while data protection refers to limitations or conditions on the processing of data relating to an identifiable individual" [2]. Personal data includes name, address, e-mail address, telephone number, IP (Internet Protocol address), MAC (Media Access Control address), GPS (Global Positioning System) location, RFID (Radio-Frequency Identification) of tags, cookies on websites, photos, video footage, personal identification number (PID), biometrical data (fingerprints, eyeshadow shooting), genetic data, education, salary, credit lending data, bank account data, health information, sexual orientation, and many other factors related to an individual whose identity is known or can be determined [3]. When the GDPR was introduced to protect this data, HEIs had to adapt to the new regulation. HEIs collect different personal data:

- Personal data about past, current, and prospective students.
- Personal data about past, current, and potential employees.
- Other personal data, related to co-workers on projects, initiatives, and other activities.

The personal data collected from students include name, address, e-mail address, high school success, courses enrolled, disciplinary offences (if any), students' logging on to HEI website, and other information. Employees' personal data include name, address, e-mail address, work experience, work results, salary, credit lending data, bank account data, teacher evaluations by students, employees' logs of on HEIs' websites, etc. Additionally, in teaching process, a special focus is put on the influence of data protection and privacy frameworks on the design of learning analytics systems [4]. Besides, authors investigated the data protection policies on higher education institutions (an example is given in [5]).

Data collected for co-workers involved in initiatives and activities include name, address, e-mail address, bank account data, and other information, depending on the type of collaboration.

Data protection is a challenge for the institutions that collect this information. If the information is not exchanged via the Internet, most institutions already have a process of keeping personal data, which the rules should adjust to a lesser extent. However, this study is particularly interested in applying the GDPR to data that are used on the Internet or stored in the cloud. As many HEIs choose this option because of the large amount of data they deal with and the more affordable cloud storage costs, they must be aware of the potential challenges. Duncan analysed articles discussing the new regulation and the possibility of its application to cloud databases [6]. The author concludes that it is best to further encrypt data whose encryption keys will not be kept in the cloud, except for additional user training. Though children and young people are particularly vulnerable, the conclusion of the round table held in Brussels (2017) is that education on the issue of personal data protection is vital for the professors themselves, and not just the students, because they do not sufficiently understand its importance. It also addressed the need to increase awareness of the period in which the data are available [7]. Additionally, improving information literacy was found to contribute to better protection of personal information.

3. Research goals and Methodology

As explained above, this paper focuses on the second section of a three-phase study. This section consists of two components:

- The first component involves conducting the questionnaire and examining how familiar students are with the GDPR just before the start of its application (April/May of 2018). The research was limited to one HEI in Croatia, which was used in the first phase of this three-phase study. The questionnaire contained 25 questions and was implemented via an online survey using the Limesurvey[1] platform, in line with principals of friendly graphic design [8]. There were two sets of the questions, with 12 questions relating to the demographic profile of participants and the remaining 13 dealing with knowledge of the concept of personal data, elements of the GDPR, and its application and mechanisms of personal data protection. The reliability of the questionnaire was examined by computing the Cronbach Alpha coefficient using the open code tool Free Statistics Software (v1.2.1) [9], [10].

- The second component consists of the identification of possible risks and problems that might result from limited knowledge about the GDPR in the student population, which can have an influence on further study. This section utilizes a risk matrix.

The main goals of this research are:

1. To identify students' attitudes in terms of knowing the basic definitions of personal data and GDPR by Croatian students (from Rijeka),
2. To analyze possible risks via risk matrix method related to misunderstanding of GDPR regulations by students.

Risk is defined as a possible uncertain situation in the future that can have a positive or negative impact. The basic idea of risk management is to anticipate the future, identify the problems that may appear, and define activities that can successfully solve these problems [11], [12]. Main risk components, in terms of risk quantification, include the following [11], [13]:

- Risk probability.
- Risk impact.

Depending on the project, data on risk components are collected by the analysis of historical data in similar situations and through various predictive simulations and models [12]. Also, experts in the problem area may be helpful in defining risk components.

In this paper, risks were identified based on questionnaire results and the analysis of experts in the field of the GDPR and education (the authors of this paper). After the risk is identified, the risk manager can utilize one of several general strategies to reduce the risk and its consequences [12], [13]:

1. Risk assumption, risk retention: This strategy is used when decision-makers decide not to take any specific action in the direction of risk resolution.
2. Risk control: In this situation, decision-makers are aware of the risk and its consequences and can solve this risk alone, so they do not require additional help. They define and implement concrete activities to deal with risk.
3. Risk transfer: In this strategy, decision-makers are aware of the existence of the risk and the strong negative impact it will have, and they look for partners to aid them in decreasing or eliminating the negative impact of the risk on the project.
4. Risk avoidance: In this strategy, decision-makers change the project's goals instead of dealing with the risk.

Depending on the risk consequences, there are several different risk types [13]:

- High risk: the consequences of the risk have a high impact on the realization and results of the project.
- Moderate (medium) risk: the consequences of the risk have a moderate impact on the project.
- Low risk: the consequences of the risk have a low impact on the project.

Table 1. Risk Matrix [11], [12], [13]

		Impact of risk on the project		
		1	2	3
The probability of risk occurrence	1	Low	Low	Moderate
	2	Low	Moderate	High
	3	Moderate	High	High

In addition to this, it is possible to classify risks more precisely using five categories (Very Low, Low, Moderate, High, and Very High [11]) or even more categories [14]. Risks are categorized into three sections in Table 1.

After a risk is identified and classified according to the Table 1., corrective strategies for each risk type have to be identified. Mostly, this is done by using the qualitative analysis with experts in some brainstorm session. It is important to determine

---

[1] http://inovacije.eu

corrective measure for each risk type for each risk, and when the risk will appear, the corrective measure that will be applied will depend on the risk type at the moment of risk appearance.

4. Questionnaire Results

The reliability of the questionnaire was verified by calculating the Cronbach Alpha coefficient. Cronbach Alpha equals 0.7891 which means that the data are reliable for using and making conclusions.

4.1. Demographic structure of respondents

In total, 141 students participated in the study, both male (N=47) and female (N = 94). This includes 80 full-time students and 61 part-time students. Many part-time students were employed, and their participation in the study required additional engagement. Full-time students spent more time at the HEI and were more willing to engage in filling in the questionnaire. In Table 2., age groups of the respondents are shown.

Table 2. Age groups of respondents

Age group	Number of respondents
18-25	102
26-35	22
36-45	15
46-55	2
>55	0

The majority of respondents (N=102) belong to the age group of 18 to 25 years; there were no participants older than 55 years in this sample. All respondents completed a four-year high school program, including 40 gymnasium and 101 vocational high schools. As the education system of the Republic of Croatia did not include computer science as an obligatory subject when respondents attended elementary or secondary school, we were interested in their previous education in this area. Six respondents did not have the opportunity to attend computer science courses during their previous education and did so only after their arrival at the HEI. The other respondents had a course in computer sciences during their previous education. The majority, 43 students, attended a course in the subject for 2 years; 29 attended it for 1 year; 27 for 4 years; 13 for 3 years; 4 for 5 to 6 years; and 5 for 8 to 9 years. When asked whether they had completed some additional form of education by attending an institution which provides that kind of service, 32 of them answered yes, while the remaining 109 respondents did not acquire knowledge in that way.

4.2. Knowledge about the GDPR

In the first question, respondents were asked to explain the term *personal data*. Responses to this question have been analysed in accordance with the definition of personal data contained in the GDPR. A qualitative analysis was carried out, since the data were in text format. The results are shown in Table 3.

Table 3. Classification of answers to the question about the definition of personal data

Response category		Number of respondents
The term *personal data*	… was not defined	17
	… was not defined, but correct example or examples are given	52
	… was not defined, but importance of personal data is recognized	9
	… was properly defined	37
	…was not properly defined	26

Most of the students (52) gave correct example(s) of personal data, but the definition of the term itself was not given; 37 students correctly defined the term; 26 gave a partial definition; and 17 students didn't define the term.

We can conclude that most respondents recognize forms of personal data, especially forms such as e-mail addresses, phone numbers, bank account numbers, photos (including any digital or similar image, such as fingerprints, corneal scans, and other body parts; i.e., biometric records), personal identification numbers, and location information. While parents' names and location information were often overlooked, a number of respondents selected them as a form of personal information. Students do not consider their grades to be personal data.

Responding to the next question in the questionnaire, participants showed knowledge of the emerging forms of personal data among those offered in Table 4.

Table 4. Knowledge about personal data forms

The form of personal data	Response	
	Yes	No
E-mail address	105	36
Phone number	118	23
Bank account	112	29
Parents' names	82	59
Personal identification number	124	17
Student grade	67	74
Photography	120	21
Data about health	80	61
GPS location	97	44

The majority of participants, 90.7%, responded positively as to whether they provided personal data, such as e-mail addresses, names, surnames, personal photos, etc., via the Internet. The results are expected, but it is somewhat surprising that the answer was not 100% positive (Figure 1.).
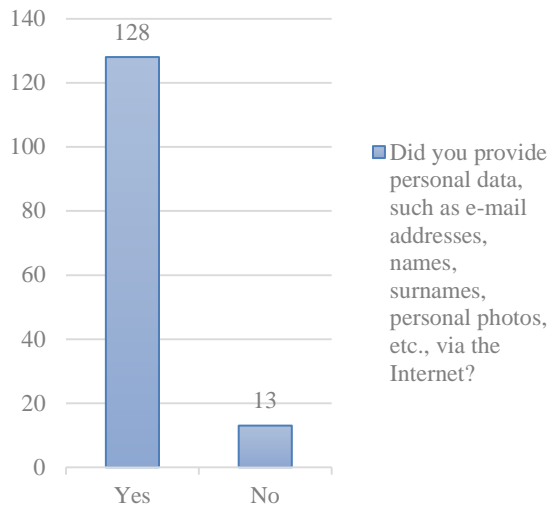

Figure 1. Providing personal data via the Internet

The period in which they provided their personal data via the internet is displayed in Figure 2.

Most of the respondents, 45, had provided personal data over the last month of the survey period. Only 7 respondents had provided their personal data during the last 6 months. 30 respondents don't know if they provided their personal data in the past.
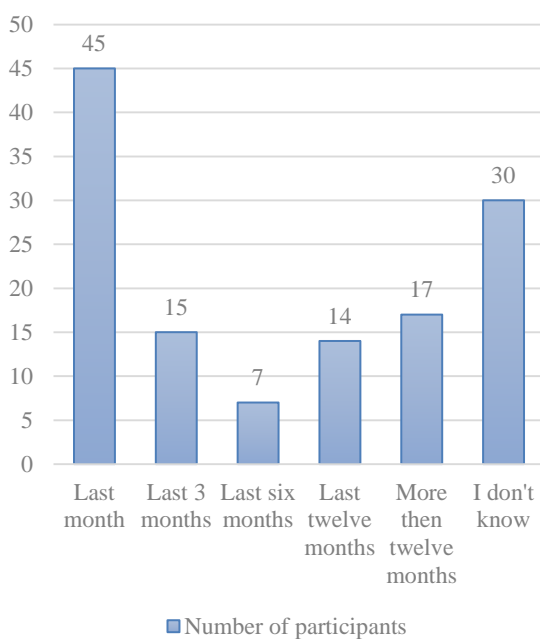

Figure 2. The period of providing personal data via the internet

Among the respondents, 79 (56.02%) were informed that someone had processed their personal data via the Internet; 30 (21.27%) were not informed; 19 (13.47%) do not know if they were informed, and 13 (9.2%) did not respond because the question was not mandatory.

When informed that their personal data would be processed, 69 (48.93%) were familiar with the purpose of collection and processing, 37 (26.24%) were not familiar with the purpose of collection and processing; 22 (15.60%) of respondents did not know the purpose; and 12 (8.5%) did not respond. Permission to use personal data was given by 67 (47.51%) respondents; 36 (25,53%) respondents did not give their permission; 25 (17.73%) didn't know if they gave their permission; and 13 did not give a response. When asked whether they are familiar with the ability to request termination of further processing of their data on a given date, respondents were almost equally divided, with 43.28% familiar and 46.26% unfamiliar, while 7 of the participants (10.44%) didn't know if they were familiar (probably they don't remember).

Figure 3. shows responses to the question, "Do you know that there is an obligation for a person/institution to enforce the protection of your personal data provided to that person/body over the Internet?" As shown, 63% of the respondents knew this, 28% did not, and 9% did not answer the question.
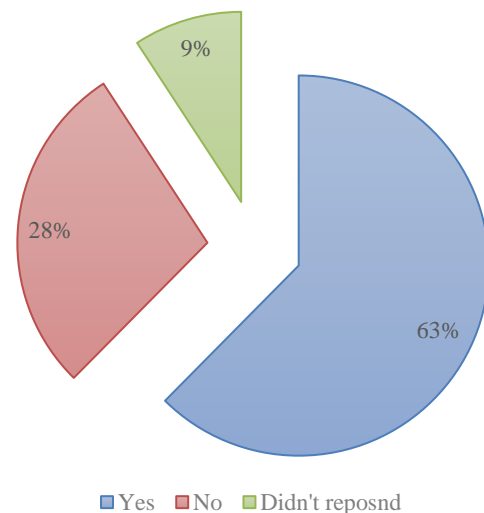

Figure 3. Knowledge about the obligation of a person/institution to enforce the protection of personal data

Considering that the majority of students knew that there is an obligation to enforce personal data protection, 94 (66.66%) participants thought about the possibility of abuse of collected personal data, 34 (24.11%) did not think about possible abuse, and 13 (9.21 %) did not respond. When asked if they had

heard about the GDPR, 60 (42.55%) respondents had, while 81 (57.44%) had not. Participants who had heard about the GDPR were asked if they knew the start date of the application. Of these, 40 (66.66%) responded that the start date is at the end of June of 2018, while the other 20 (33.33%) did not know when the application started. Respondents who knew about the GDPR gave the answers below in Table 5.

Table 5. Action to improve the protection of personal data

Action to improve the protection of personal data by decree	Yes	No
I must give explicit permission to collect, process, and publish my personal data.	46	14
I will get information in a clear and understandable way how and for what purpose my data is processed.	42	18
At any time, I may request a correction or amendment of my personal data.	31	29
I have the option to request that the companies and organizations that process my personal data delete it.	48	12

5. Student Problems Related to the GDPR and Risk Matrix Analysis

The general regulation for individual protection, in terms of personal data processing, strongly relates to students. Their data are being collected and used by HEIs from the date of enrolment until the completion of their studies, and even later. This includes names, student ID numbers, e-mail addresses, dates of birth, photos, telephone numbers, and other information. HEIs are responsible for informing their students:

- Which data are collected
- What is the basis for collecting and processing the data (legal obligation, consent of the data owner, and other legitimate reasons)
- How those data are collected
- The purpose of data collection and how the information will be used and processed
- How the data will be stored
- How long the data will be stored
- Who is able to access the data
- What are the students' rights in terms of data collecting
- When are certain data deleted or destroyed

In this context, it is important to determine how well students know their rights and how that awareness can be increased.

The students' misunderstanding or ignorance of the basic concepts of the GDPR introduces two fundamental issues:

1. Students aren't aware of the use of their personal data by others; they don't know who can access them, whether they are used for unwanted purposes, whether they are available to other students and the public, if they are used for promotional activities of the HEI, etc.

2. Students aren't aware that other students' data are protected; they may want to see the tests results of other students, their own tests results after the expiration of test keeping date, other students' activity data, like how certain students voted on some issue, etc.

Both components can be analysed by students and HEIs alike. In the process of adapting to the GDPR, HEIs introduced a rulebook concerning protection, the processing of personal data, and privacy policies. By introducing those documents and enforcing their application, HEIs ensure that personal data are protected according to the GDPR. However, this does not mean that students do not have to know the regulation. Understanding the GDPR enables data protection reaction, if it is not implemented by HEIs, and also the prevention of some malicious acts.

Table 6. Risk analysis

Risks	Type	Strategies
The data are used by unauthorized persons.	L	No action.
	M	Education about the GDPR.
	H	Reaction in terms of legal actions.
The data are available to other students and public.	L	Education about the GDPR.
	M	Reaction in terms of legal actions.
	H	
Students want to see the tests results of other students.	L	No action.
	M	Education about the GDPR.
	H	Education about the GDPR.
Students want to see their tests results after the expiration of test keeping date.	L	No action.
	M	Education about the GDPR.
	H	Education about the GDPR.
Students came to protected data in an illegal way.	L	Education about the GDPR
	M	Increase security levels of all systems.
	H	

In Table 6., there is an analysis of possible risk situations and defined risk management strategies. The management (corrective) strategies are identified by qualitative analysis of field experts (teachers and

members of HEI management board). The strategies depend on the risk level (type) of a certain risk. The strategies will go into action at the time of risk appearance depending on the level of the risk at the time of appearance.

The most logical strategy, as shown in Table 6., is implementing education about the GDPR. Hence, the main conclusion of this analysis is to recommend that the management of HEIs organize instruction and workshops related to this topic.

6. Conclusion

The GDPR introduces some novelties in personal data handling and management. This requires the adaption of both organizations and individuals. It is important to know the basic concepts of personal data and the elements of the GDPR.

This study was focused on higher education institutions and students. The results of the questionnaire show that there are still some issues related to understanding the basic concepts of the GDPR among student populations. The risk matrix analysis concluded that the best strategy, in this case, would be to organize workshops and lectures related to the GDPR and personal data. That can decrease possible problems involving failure to comply with the provisions of the GDPR.

This research involves the second part of a three-phase study related to students' understanding of the GDPR. The third part of the study is planned to be implemented in April of 2019.

Acknowledgements

References

[1] Vukobrat, N. B., Marković, M. G., & Debeljak, S. (2017, January). Do we know how to protect personal data in the virtual environment?. In *4th International Multidisciplinary Scientific Conference on Social Sciences and Arts SGEM 2017*.

[2] Politou, E., Alepis, E., & Patsakis, C. (2018). Forgetting personal data and revoking consent under the GDPR: Challenges and proposed solutions. *Journal of Cybersecurity*, *4*(1), 1–20.

[3] Span d.o.o. (2018). Što je GDPR – GDPR2018. Retrieved from: https://gdpr2018.eu/sto-je-gdpr/ [accessed: 21 July 2018].

[4] Hoel, T., Griffiths, D., & Chen, W. (2017, March). The influence of data protection and privacy frameworks on the design of learning analytics systems. In *Proceedings of the seventh international learning analytics & knowledge conference* (pp. 243-252). ACM.

[5] Frank, R., & Wagner, L. (2018). Understanding the Importance of FERPA & Data Protection in Higher Education. An Application: Website at La Salle University. Retrieved from: https://digitalcommons.lasalle.edu/mathcompcapstones/36/ [accessed: 10 December 2018].

[6] Duncan, B. (2018). Can EU General Data Protection Regulation Compliance be Achieved When Using Cloud Computing. In B. Duncan, Y. W. Lee, & A. Olmsted (Eds.), *CLOUD COMPUTING 2018 : The Ninth International Conference on Cloud Computing, GRIDs, and Virtualization* (pp. 1-6). [28010] (Cloud Computing 2018). IARIA.

[7] Milkaite, I., Verdoodt, V., Martens, H., & Lievens, E. (2017). The General Data Protection Regulation and children's rights: questions and answers for legislators, DPAs, industry, education, stakeholders and civil society. Roundtable Report.

[8] Couper, M. P., Traugott, M. W., & Lamias, M. J. (2001). Web survey design and administration. *Public opinion quarterly*, *65*(2), 230-253.

[9] Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *psychometrika*, *16*(3), 297-334.

[10] Wessa, P. (2017). Cronbach alpha (v1. 0.5) in Free Statistics Software (v1. 2.1). *Office for Research Development and Education*, Retrieved from: https://www. wessa. net/rwasp_cronbach. wasp. [accessed: 05 September 2018].

[11] Sikavica, P., Hernaus, T., Begičević Ređep, N., & Hunjak, T. (2014). *Poslovno odlučivanje*. Školska knjiga Zagreb.

[12] Kadoic, N., & Kedmenec, I. (2018, May). ERA metamodel of the analytical hierarchy process and risk matrix. In *2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)* (pp. 1322-1327). IEEE.

[13] Divjak, B., Spahić, A., Brodar, K., Stapić, Z., Orehovački, T., Munđar, D., … Lovrenčić, S. (2009). *Projektni ciklusi u znanosti i razvoju*. (B. Divjak, Ed.). TIVA tiskara, FOI Varaždin.

[14] Ruge, B. (2004). Risk matrix as tool for risk assessment in the chemical process industries. In *Probabilistic Safety Assessment and Management* (pp. 2693-2698). Springer, London.