Proceedings of the 11th International Conference on Applied Informatics
Eger, Hungary, January 29–31, 2020, published at http://ceur-ws.org

The Level of Information Security Awareness of First-Year University Students

Gábor Törley

Eötvös Loránd University, Faculty of Informatics, Budapest, Hungary
gabor.torley@inf.elte.hu

Abstract

According to the results of a representative survey by ESET Hungary Ltd. and statistics by Eurostat, in Hungary more than one million users visit infected webpages despite of the warnings of their antivirus program and almost every second individual caught a virus or other computer infection (worm, Trojan horse, etc.). These data are similar in Slovenia, in Croatia, in Slovakia and in Bulgaria. This can be caused by the low level of security awareness.

According to the first International Computer and Information Literacy Study (ICILS), understanding of online safety and security issues are part of the definition of computer and information literacy. In 2012, the PISA assessment results show that among countries with deteriorating performance in digital reading, Hungary was one of the countries what shows the biggest declines in performance among their weakest students.

This study discusses three topics:

(1) What are pupils taught on e-safety, privacy and information security in Hungary and how much lesson hours can teachers use for these topics. This part of the study shows how solid is the "basement" of security awareness knowledge of an average pupil.

(2) What level of information security awareness can be expected from an aver-age first-year university student from different fields of knowledge without any university level teaching? A questionnaire on important concepts and user behavior (password policy, social networks, etc.) can answer this question.

(3) How and what can we teach these people in university in order to strengthen their awareness? This is an important question because most of these students will manage other people's personal data at their workplaces,

but how could they manage them securely if they cannot be vigilant with their own personal data.

Keywords: education, information security awareness, empirical study, teaching methods

MSC: 97Q99

1. Introduction

Computer and information literacy (CIL) is defined as "an individual's ability to use computers to investigate, create, and communicate in order to participate effectively at home, at school, in the workplace, and in society". [4]

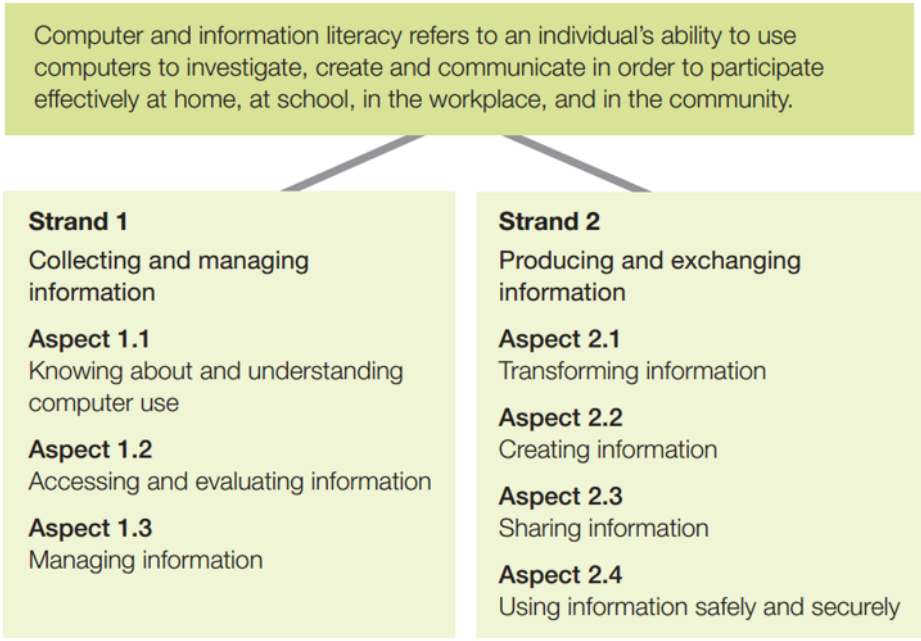

Figure 1: Conceptual structure of the CIL framework [4]

Aspect 2.3 and 2.4 (Fig. 1) have information security issues, especially Aspect 2.4. Sharing information is about a person's ability of using email, wikis, blogs, instant messaging, sharing media, and social networking websites. Internet-based communication platforms provide a lot of possibility for users to share information. With this facility comes the potential for misuse, especially when dealing with personal information. Using information safely and securely also contains risk identification, prevention, the parameters of appropriate conduct and responsibility of users to maintain a certain level of technical computer security, for example using

strong passwords, keeping virus software up to date, and not submitting private information to unknown publishers.

According to the PISA Assessment results (2012), among countries with deteriorating performance in digital reading, Hungary was one of the countries what shows the biggest declines in performance among their weakest students. 32% of the Hungarian students performed low in digital reading. [10] Understanding of online safety and security issues are part of digital reading. [4]

The results of PISA 2018 assessment [11] have shown the same: the Hungarian students have performed below the OECD average in reading. Over 25% of Hungarian students performed below Level 2 proficiency in reading. At Level 2, students can identify the main idea in a piece of text of moderate length. Evaluating and reflecting has always been a part of reading literacy. In the era of digital reading, readers are now confronted with ever-growing amounts of information, and must be able to distinguish between what is trustworthy and what is not. This is also a part of information security awareness.

According to the results of a representative survey in 2011 by ESET Hungary Ltd., more than 1 million Hungarian Internet users open dangerous webpages, despite of warnings from their antivirus system. Moreover, 10% of Hungarian Internet users switch off intentionally their antivirus software in response. This is mostly typical in the age group 18-29, that is 17% of the Hungarian Internet users, which is especially disquieting, because this group includes those young men and women who just finished their secondary education. [2] These results are not surprising. Low level of digital reading can cause low level of information security awareness (ISA).

Security experts generally agree that people (the human factor) are the greatest source of information technology (IT) security-related problems. Statistics showed that the majority of security issues are caused by insiders, and the caused damage can be more serious than anything elaborated by hackers from outside. [12]

In my empirical and exploratory research, I show the average ISA level of first-year students at National University of Public Service, Faculty of Political Sciences and Public Administration and Eötvös Loránd University, Faculty of Informatics, Faculty of Science and Faculty of Social Sciences, I find out the reasons and correspondences behind the results, finally I suggest solution to strengthen the ISA level, focusing on secondary and B.A. education as well as workplaces in Public Administration and IT business.

2. Literature review

Livingstone et al (2011) [9] lead a survey which investigated key online risks: pornography, bullying, receiving sexual messages, contact with people not known face-to-face, offline meetings with online contacts, potentially harmful user-generated content and personal data misuse. They showed that younger children tend to lack skills and confidence. However, most 11-16 year olds can block messages from those they do not wish to contact or find safety advice online. Around half can

change privacy settings on a social networking profile compare websites to judge their quality or block spam. They claim that digital skills training is needed in order to ensure that all children reach a minimum basic standard and to prevent digitally isolated and unskilled children.

According to van der Walt et al (2008) [13] and Kruger et al (2010) [8], one's information security vocabulary come from the set of familiar words related to information security. Such a vocabulary will develop over time and with that, a person can communicate and acquire new knowledge. Based on this approach, Kruger et al (2010) [8] developed a questionnaire, which consists of two sections – a first section to perform a vocabulary test and a second one to evaluate respondents' behaviour. The results showed that there is a connection between the knowledge of concepts (vocabulary) and the behaviour as well as that a vocabulary test can support to identify specific areas for security education.

Krasznay and Törley (2015) [7] showed a short overview on Hungarian secondary education. They found that sample curriculum (developed by major textbook publishers) use only around 4-5% of informatics lesson hours for data protection and information security in grades 5-12. This number and portion of lesson hours are very low.

3. Methodology

A questionnaire with 41 questions was developed in February 2016. and February 2017. It contains 5 question-groups: simple statements on ISA, questions on secondary school studies, theoretical, practical, demographic questions. The target audience is first- year students from National University of Public Service, Faculty of Political Sciences and Public Administration and Eötvös Loránd University, Faculty of Informatics, Faculty of Science and Faculty of Social Sciences. They have never met with any ISA training/course before, so that is why this questionnaire can measure these students' input knowledge on privacy, information security and data protection.

Main goal of the questionnaire is to identify those topics which were uncovered during high school studies, to find the connection between the theoretical and practical answers as well as the high school studies and to measure how well the students predict their level of ISA (i.e. the connection between the "ISA statements" and the overall results of the questionnaire).

4. Results

627 first-year students completed the questionnaire (67% of all first-year students), 58% of them are male and 42%-of them are female. 82% of them are 18-20 years old.

There were two simple statements, so called "ISA statements": (1) I am a

security-aware person, (2) I am fully aware of the basic concepts related to information security awareness (see Fig. 2).

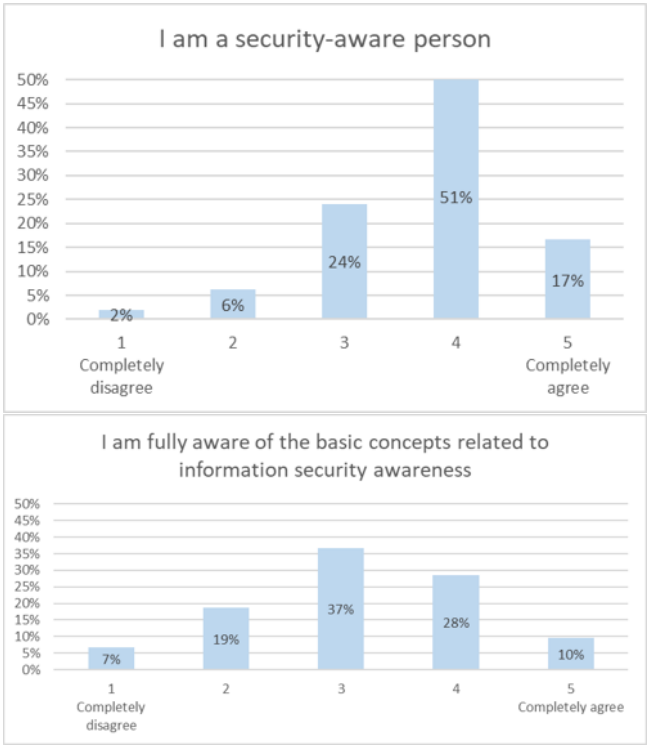

Figure 2: Distribution of the answers to the "ISA statements"

The great majority of the students (68%) think that they are security-aware but a notable proportion of them (62%) do not really know or are unsure what the basic concepts are on ISA. This can mean, they have experiences rather than real knowledge on ISA. This uncertainty of theoretical knowledge impacted the average percentage of the answers to the theoretical questions, which was only 39%, unlike the average results of the answers to the practical questions, which was 66%.

Comparing the results between the students who think themselves security aware and who do not (with Ordinary Least Squares (OLS) test), I found that students with higher evaluation performed significantly better on practical questions (coefficient=0.511, p=0.003<0.05, ceteris paribus[1] and reached significantly more points on the practical and theoretical questions together (coefficient=0.654, p=0.002<0.05, ceteris paribus). They do not perform significantly better on other topics (social network, security of smartphones).

[1] *Ceteris paribus* or *caeteris paribus* is a Latin phrase, literally translated as "with other things the same," or "all other things being equal or held constant.")

Those students who think themselves sure in basic ISA concepts (38% of them) reached significantly more points on the practical and theoretical questions together than other students (coefficient=0.712, p=0.0005<0.05, ceteris paribus).

There were 7 theoretical questions on basic definitions on ISA (based on the Hungarian secondary school curriculum framework) and 17 questions on how students would behave in practical situations (privacy issues on social networks, secure use of smartphones, passwords, wireless networks, computers and Internet browsers as well as the practical use of privacy). Each correct answer is worth one point. The scoring system was the following (Table 1):

Grade	**Bottom point limit's ratio**
1(worst)	0%
2	50%
3	62%
4	75%
5	88%

Table 1: Scoring system

If this questionnaire would have been a real test, 22% of the students would have failed, 41% of them would have got grade 2, 33% of them would have got grade 3 and 4% of them would have got grade 4. There would not have been students who would have got grade 5 (see Fig.3).

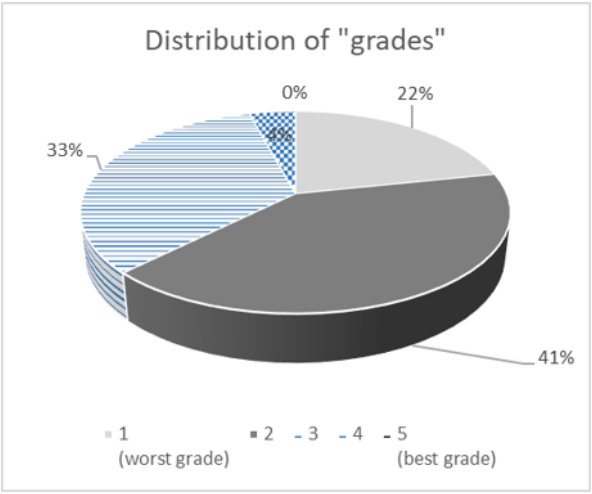

Figure 3: Distribution of "grades"

More than 78% of the students share their personal data and posts only to their friends, to self-created closed groups or to only themselves. Contrary to this, 28% of the students share special (sensitive) data (religion, sexual orientation, political

views) with their social network. According to the European Union's General Data Protection Regulation (GDPR) [6], special data shall mean "personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation"[2]. This kind of data can be processed if "the data subject has given explicit consent to the processing of those personal data for one or more specified purposes" or when processing is necessary. For example: workplaces or medical institutions need to process health data; at a serious accident if the data subject is physically or legally incapable of giving consent then in order to protect the vital interests of the data subject, data processing is needed; processing is necessary for reasons of substantial public interest, etc.[3] More than one-fourth of the students "throw away" this defense of the law because they do not know or do not care about what special (sensitive) data are.

Only 46% of the students use their password securely i.e. it contains lower and uppercase letters and numbers; it is not a regular word; it does not refer their personal data and its length is more than 8 character. 29% of these students (only 11% of the whole sample) use these principles regarding password hints. This means that 89% of the students think of password hints as a support instead of handling it as password of password. 21% of passwords and 38% of password hints refer to personal data. Great majority of students use easy-to-hack password to their E-mail account.

25% of the students do not use any security setting to access their smartphones. 20% of the sample think that "moving the lock on the screen" is a security setting which is not true. Android smartphones work this way when a user leaves security settings on default values.

According to Fig. 4, every third student did not have any lesson hours on information security and data protection in secondary school. More than a half of the students (57,2%) had one or less lesson hour on these topics, which is very few. This can also explain why students are unsure on their theoretical knowledge (see Fig. 2 above).

Computer science and informatics develop fast that is why the validity of knowledge decreases during the years without learning or practice. Fig. 5 shows that 60% of students had their last informatics lesson since 0 or 1 year ago. Comparing the results (with Ordinary Least Squares test), those students who had their last informatics lesson since 0 or 1 year ago, did perform significantly better on practical questions (coefficient=0.335, $p=0.038<0.05$, ceteris paribus) and on the overall result (coefficient=0.435, $p=0.03<0.05$, ceteris paribus).

Eurodyce (2011) [3] defines the main topics on data protection and information security: Online safe behaviour, Privacy issues, Cyberbullying, Downloading issues, Safe use of mobile phones, Contact with strangers, Safe use of social networks, Use

[2]GDPR Article 9. Paragraph 1.
[3]GDPR Article 9. Paragraph 2. (a)-(j)

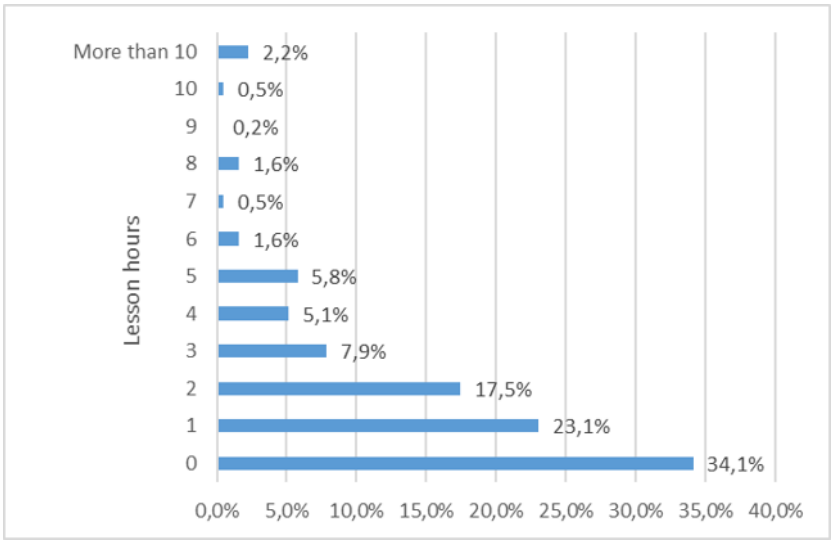

Figure 4: Distribution of answers to question "How much lesson hours did you have (alltogether) on data protection and information security at grades 9-12."

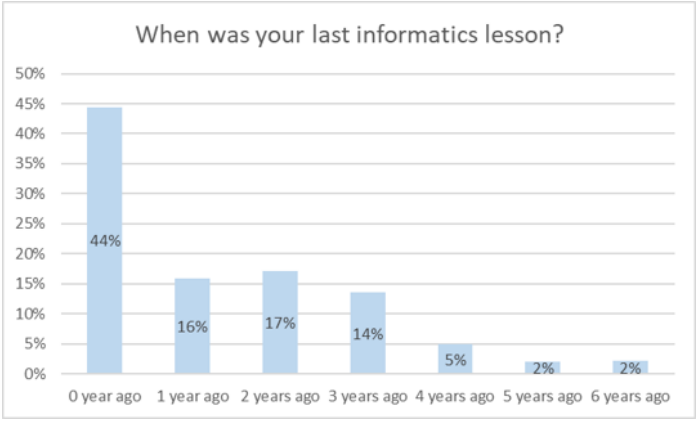

Figure 5: Distribution of when students had their last informatics lesson

of antivirus softwares, Password policy. I asked the students which topic(s) they did learn. Fig. 6 shows the distribution of the answers.

The values are very low, especially "Safe use if mobile phones". 42% of students taught on password policy and the same amount, 46% of them use their password securely. I did not find any connection between the answers above and the overall results.

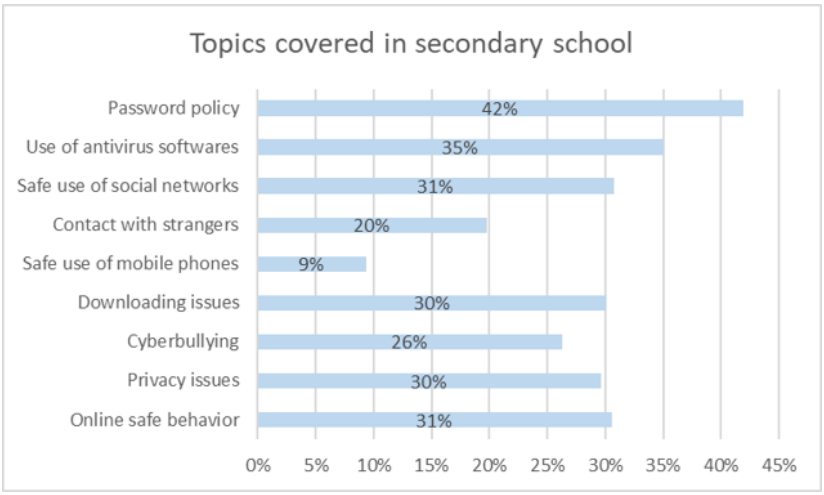

Figure 6: Distribution of which topics were covered in secondary school

Students have inhomogeneous knowledge on these topics. 54% of them were thought on two or less topics from the main topics of [3]. This can be caused by the unbalanced local curriculum of the informatics subject in high school. This can mean that every second student do not have a solid foundation on information security and data protection. That is why their training should be started from the beginning, from the basic concepts of ISA at university level.

There was a question on wireless network security in the questionnaire. I asked the students which the recommended security protocol is for a wireless network at home. Only 46% of them knew the correct answer (see Fig. 7).

Almost 17% of the students chose insecure protocols, and almost one-third of them chose protocols which do not exist . This can mean that students do not have appropriate theoretical knowledge on wireless network security and they usually do not check their wireless settings on their smartphones/computers/wireless routers, they "just use" it as a service. this lack of knowledge can include the possibility of deceit.

In Hungary, high school pupils can choose informatics as a secondary school-leaving exam subject. They can choose the level of the exam, too: standard or high level. According to the answers, 27% of the students took standard level exam on informatics, and 18% of them took high level exam. These students may have positive attitude and/or talent to the subject. Comparing the results (with Ordinary Least Squares test), those students who took school-leaving exam on informatics (at any level), did perform significantly better on the practical and theoretical questions together (coefficient=1.17, $p=0.00000000199<0.05$, ceteris paribus) and on theoretical (coefficient=0.332, $p=0.0005<0.05$, ceteris paribus) and practical questions as well (coefficient=0.844, $p=0.0000000999<0.05$, ceteris paribus). They

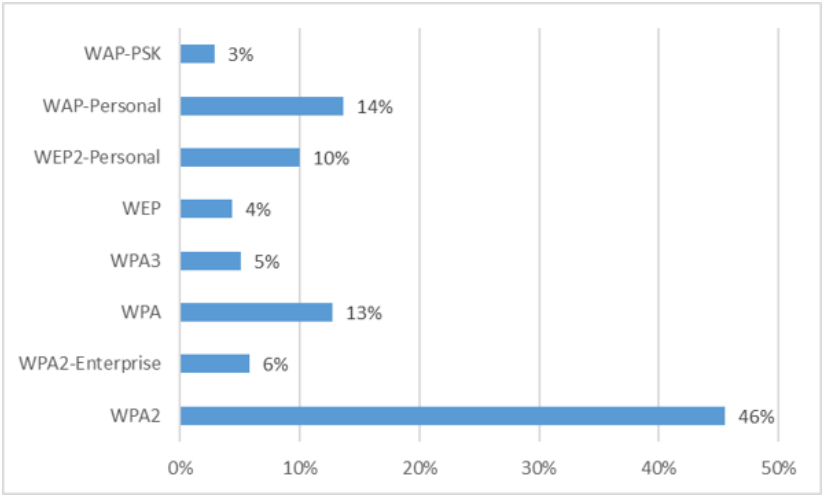

Figure 7: Distribution of answers on question "Which is the recommended security protocol for a wireless network at home?"

needed to have more lesson hours (because of the exam), so this result refers to Fig. 5 Students' preparation (and extra lesson hours) impacted more their theoretical and practical knowledge.

5. Conclusions

In secondary school and university curriculums and textbooks, the definition of data protection and protection of data should be separated more obviously because these definitions were mixed up by the students. Data protection means "legal processing of personal data; principles, rules, procedures, tools for data processing and methods which guarantee the protection of data subjects" [5], protection of data means those protection methods which are executed on data in the interest of the data subjects and the data processors. The "leading character" of data protection is the data subject, protection of data focuses on the data. That is why, the concept of information security can approach from two sides: from legal and from technical side as well as there is one goal of these two sides: data of data subject should be protected. This is what I call "unified approach" which I suggest following in secondary school and university level, i.e. to teach data protection and information security in the same course/subject.

Since information security awareness is the ability of recognize or avoid behaviors that would compromise information security, that is why teaching methods cannot base on only lexical knowledge however getting to know the appropriate vocabulary is important. Someone's behavior can be changed easier by experiences that is why experimental learning seems a working teaching method in order to

strengthen pupils' ISA.

The results of the questionnaire showed that a first-year student's input ISA knowledge and level is low. That is why universities' role is important, especially those which trains new employees for Public Administration and for IT business to increase this level. Since the validity of ISA knowledge decreases fast, workplaces should pay attention for regular yearly training for their employees on ISA. Usually Public managers often use reactive ways to resolve information security problems. Reactive and preventive ways should be used together and among the preventive ways are education and training. [1]

Other important conclusion of the questionnaire that a balanced curriculum is needed in secondary and higher education as well as on trainings at workplaces. Unbalanced curriculum can cause "white patches" in ISA knowledge and this study showed what are those topics which should be discussed at secondary and higher education and on workplace trainings.

There are several workplaces where employees manage other people's personal data, but how could they manage them securely if they cannot be vigilant with their own personal data. I believe that security awareness is a way of thinking which is teachable and learnable. As several studies, cases and international experience show lack of human awareness, knowledge and concrete skills quite often the key causes of hacker attacks, violations, data compromise, or system breakdowns. Every workplace has or should have information security policy, but without enough information security awareness, every policy is useless.

This study showed that those students who took secondary school-leaving exam on informatics performed significantly better. That is why if informatics could be a compulsory subject then first-year students' level of ISA could be higher.

Acknowledgements. I would like to thank to VSL Ltd. for letting me use the Evasys system for free.

References

[1] D'Arcy, J., Hovav, A. Deterring internal information systems misuse, *Communications of the ACM* 50(10), 113-117 (2007)

[2] ESET Hungary Ltd.'s survey 2011., `http://www.eset.hu/hirek/kivancsisagunk_fertoz?back=/hirarchivum\%3Fpage\%3D9` (in Hungarian) last accessed April 6. 2016.

[3] Eurodyce 2011. Key Data on Learning and Innovation through ICT at School in Europe 2011, *European Commission, Education, Audiovisual and Culture Executive Agency,* ISBN-978-9-2920-1184-0, (2011), `https://publications.europa.eu/en/publication-detail/-/publication/8f864668-0211-4a40-bc14-65bf1a97b6a8` last accessed Januar 24. 2020.

[4] Fraillon, J., Schulz, W. and Ainley, J. International Computer and Information Literacy Study assessment framework, *Amsterdam, the Netherlands: Inter-*

national Association for the Evaluation of Educational Achievement (IEA), https://www.acer.org/files/ICILS_2013_Framework.pdf Last accessed Januar 24. 2020.

[5] National Authority for Data Protection and Freedom of Information Defining vocabulary of data protection (translated from Hungarian) - http://www.naih.hu/adatvedelmi-szotar.html - last accessed Januar 26. 2020.

[6] Regulation (EU) 2016/679 of the European Parliament and of The Council

[7] Krasznay, Cs., Törley, G. E-safety, privacy and information security: Requirements in Public Administration, *In: Alexander B., Golob, B., Hansen, H., Kőnig, B., Müller-Török, R., Prosser A. (eds.) Central and Eastern European eGov Days 2015: Indepedence Day: Time for a European Internet?* pp. 431-441, Austrian Computer Society, Vienna, Austria (2015) ISBN 978-3-85403-308-0

[8] Kruger, H. A., Drevin, L. and Steyn, T. A vocabulary test to assess information security awareness, *Information Management and Computer Security* 18(5), 316-327 (2010)

[9] Livingstone, S., Haddon, L., Görzig, A. and Ólafsson, K. Risks and safety on the internet: the perspective of European children: summary. EU Kids Online, Deliverable D4, *EU Kids Online Network, London, UK.* (2011)

[10] OECD 2015. Main Results from the PISA 2012 Computer-Based Assessments, in Students, Computers and Learning: Making the Connection, *OECD Publishing, Paris*

[11] OECD 2019. PISA 2018 Results (Volume I): What Students Know and Can Do, *PISA, OECD Publishing, Paris* https://doi.org/10.1787/5f07c754-en

[12] Pescatore, J. High-Profile Thefts Show Insiders Do the Most Damage, *Gartner First Take* (2002), https://www.gartner.com/doc/379171/highprofile-thefts-insiders-damage - Retrieved: April 6. 2016.

[13] Van der Walt, M., Maree, K., Ellis, S. A mathematics vocabulary questionnaire for use in the intermediate phase, *South African Journal of Education,* 28 489-504 (2008)