DOI: https://doi.org/10.48009/2_iis_2023_102

A quantitative study on the usage of a cryptographic software tool for data and communications encryption

Mel Tomeo, *Purdue Global University, mel.tomeo@purdueglobal.edu*
Wilfred Mutale, *Duquesne University, mutalew@duq.edu*
Matthew Kisow, *Community College of Allegheny County, mkisow@ccac.edu*
John J. Scarpino, *Robert Morris University, scarpino@rmu.edu*
Vasilka Chergarova, *Florida International University, vchergar@fiu.edu*

Abstract

The purpose of this study was to provide a comprehensive understanding of the usage of software encryption and its importance in securing data and communications. Although previous studies about data and communication encryption have focused on cybersecurity, very few have given attention to examining the relationship between the use of cryptographic software with data and communication encryption. Software encryption tools can enhance data privacy and assist in protecting sensitive data. To fill this research gap, this study used a quantitative study on the usage of a cryptographic software tool for data communication encryption. Data was collected through a survey deployed on LinkedIn between April 1, 2023, and April 29, 2023. The target population for the survey was college instructors and working professionals in the field of IT. A total of 61 participants completed the survey. The findings indicated that a stronger cryptographic tool and data encryption would increase users' perception of data security when sending and receiving sensitive data. This study suggested that stronger cryptography and data encryption would increase users' confidence when sending and receiving sensitive data.

Keywords: **e**ncryption, **e**ncrypted messages, **c**ryptographic tools, data privacy, data security, data encryption, data sovereignty, GDPR

Introduction

Cybersecurity threats are increasing year after year and are not slowing down. Cyberattacks are becoming more sophisticated to gain access to individuals' sensitive and personal information. It is now an absolute need to encrypt data to keep organizations' information and systems protected.

Encrypting data is one of the most effective tools available to keep your important information out of the hands of unauthorized individuals (Dizon, 2023). Encryption is the process through which data is converted and hidden so that it is harder but not impossible to be accessible to unauthorized users (Dechand et al., 2019). The main purpose of encrypting data is to help protect companies' and individuals' sensitive information. There are large amounts of sensitive data that are currently managed and stored online (Papoutsakis et al., 2021). Companies store their information in either the cloud or on connected servers. Encrypting sensitive data is one of the first layers to protecting your data and staying cyber resilient. Software encryption tools are one of the most helpful tools that are being used in cybersecurity to defend against brute force, malware and ransomware attacks, and cybersecurity attacks.

Using an encryption software tool to protect private and sensitive information can enhance the security of communication between applications and web/database servers (Johnson et al., 2018). Even if an unauthorized person gains access to sensitive data, if the data is encrypted, then they will not be able to read and decipher the data without access to the encryption key. It is very important to encrypt data when involving personal and sensitive data on the cloud and computer systems due to the consequences and ramifications that it can have if the private information was able to be read by an unauthorized individual (Reuter et al., 2021). As organizations start to move to hybrid and multi-cloud environments, organization concerns are starting to grow about their cloud security and how they will protect data across complex environments (Bajaj et al., 2022).

Literature Review

The procedure to encrypt data is to take an encryption key and use an algorithm to convert the information or readable data into unreadable data (Shen, 2021). Encryption can be explained by having a system of mathematical algorithms that encodes the user's data so that only the presumed recipient should be able to read and understand it (Anugurala & Chopra, 2016). To decrypt a message, the recipient would need to have the correct decryption key to decode the unreadable data back into readable information. In the following sections, the advantages and disadvantages of encrypting data will be discussed.

Advantages of Encrypting Data

An advantage to using data encryption is that it allows the user's data to remain separate from the location where the device’s security is stored. This provides one extra layer of security between the device and the data. Although not 100% guaranteed, data encryption does circumvent potential complications when data breaches occur. Data breaches do occur and by encrypting your data, you improve your chances of having the data readable. An advantage to encrypting data is that regardless of how the information is transferred, the data will still be secured and only viewed by the individual that has the correct decryption key to decode it. For example, if the information were sent via email but not to the correct person, it would not be viewable because it was encrypted.

Another advantage to encrypting your data is that it helps create a sense of confidentiality. According to a survey developed by CIGI-Ipsos Global Survey, about half of the participants said they were more concerned about online privacy now than a year ago. With the increasing data breaches that are occurring every year, knowing that you are encrypting your data can help create a sense of trust when sharing personal data with an individual or business.

Disadvantages of Encrypting Data

A disadvantage to using data encryption is that if you use it on all your files in your system, being able to access and modify the files daily could become a hassle (Kang & Deng, 2023). Having to continuously decrypt files on a file system where you might be constantly creating, moving, modifying, copying, and sharing data with others could potentially create a work environment that is difficult to manage. Managing and keeping the files organized regarding what files are encrypted and shared with the correct individual could also start to become time-consuming. Having different keys for different files can easily become a problem if not managed correctly.

Another disadvantage to using data encryption technology is how it integrates with other existing programs and applications in the network (Kumar et al., 2010). Testing of the data encryption technology would need to occur before implantation since it can negatively impact routine operations within the system. A disadvantage to using data encryption technology could be the cost. Depending on the size of the file system, this technology can require a lot of resources like data processing, time consumption, usage of various algorithms for encryption, and decryption.

Research Methodology

Research Question and Hypotheses

The main research questions that this study addressed were:

RQ1*: Is it important to encrypt messages, information, and communication using cryptographic software?*

RQ2: *What is the level of difficulty in using a cryptographic software tool to send encrypted messages?*

The following hypotheses were tested:

H_1. *The participants will perceive that sending encrypted messages using modern operating systems will find the command-line interface is archaic and hard to use*.

H_2. *The participants will perceive that sending encrypted messages using a cryptographic software tool will find that it is less intuitive than the command-line interface.*

H_3. *The participants will perceive that encrypting messages and data is necessary*.

H_4. *The participants will perceive that all data and communications in every application should be encrypted.*

H_5. *The participants will perceive that encrypting data and communication should be a top priority when developing an application.*

Questionnaire Development and Testing

The questionnaire contained ten questions and were developed based on prior literature. A demographic question was asked to gather the age range of the participants. A pretest was conducted on a group of participants who completed the questionnaire by themselves, without intervention or support from the researcher. The pretest was given to participants from a specific targeted audience. The purpose of the pretest was to validate the questions on the questionnaire. The questionnaire was converted into a survey and created through a website (https://www.google.com/forms/about/).

The researchers decided to use Google Forms because of its reputation for stability and for the simple appearance of the interface that it provided. Google Forms uses traditional web widgets such as checkboxes and radio buttons. This interface helped reduce the number of instructions on how to reply to the questions. Google Forms was chosen by the researchers due to the built-in functions to analyze the results of the data collection. These tools have been tested and validated by previous studies. The tools that were provided by Google Forms were at no cost to the participants or researchers.

Data Collection Methodology

The usage of cryptographic software was measured with previously validated instruments. The instruments used in the study were a series of survey questions that were measured on a 5-point Likert-type scale in which 1 denoted “Strongly Agree (SA),” 2 denoted “Agree (A),” 3 denoted “Neither Agree nor Disagree (NAND),” 4 denoted “Disagree (D),” and 5 denoted “Strongly Disagree (SD).”

The participants for the survey were sent a link through email or got access to the questionnaire through a LinkedIn post. The participants were able to access the survey between April 1st, 2023, to April 29th, 2023. Participants were given an introduction and the purpose of the survey before being asked to take it. Participants were expected to fully understand the purpose of the survey and agree to the terms and conditions before proceeding to complete the survey. The targeted participants were individuals who had previously worked or taught in software security.

The purpose of the survey was to collect data and analyze the results to add to the body of literature regarding the importance of using a software security tool to send encrypted messages. Surveys and questionnaires are widely used in research to target a specific population with questions designed to measure and collect data about a specific topic (Alvarado et al., 2016). This technique provides precise calculations of the variables that are being used in the study.

The following statements in the questionnaire were given to the participants of this study:

RQ_1: Please select your age range.
RQ_2: Please select your gender.
RQ_3: Please select your job status.
RQ_4: I believe that sending encrypted messages using a cryptographic Software tool would be easy to use and understand.
RQ_5: I believe that sending encrypted messages using a cryptographic Software tool would not be easy to use and understand.
RQ_6: I believe that all data and communication in every application should be encrypted.
RQ_7: I believe that sending encrypted messages using the command-line interface would be easy to use and understand.
RQ_8: I believe that sending encrypted messages using the command-line interface would not be easy to use and understand.
RQ_9: I believe that all data regarding private information should be confidential and should be a top priority when developing an application.
RQ_{10}: I believe that all private information should be stored on a server that is stored in the country where I reside.

Questionnaire Distribution

The approach to invite participants to the survey was done online through LinkedIn. This was the only contact with the participants, and it explained the purpose of the research, who the researchers were, and the average time that would be spent to complete the questionnaire. Google Forms provided a header for the survey questionnaire to include additional information for the participants. This helped show the participants that the research was focused on a specific topic. The Institutional Review Board for Protection of Human Subjects in Research (IRB) approval was obtained prior to recruitment of subjects.

Data Analysis

To ensure the quality of our data collection, a thorough review of the collected questionnaires, removing incomplete responses and inconsistent answers. After this screening process, the partial least-squares structured equation modeling (SmartPLS3) was used to analyze a total of sixty-one questionnaires. This method is recommended for testing complex and less established theories, and this method was used to create a theoretical framework to visualize our variables and hypotheses. The results, including the descriptive statistical analysis of the sample data, are presented in the following section.

Results

Demographics

The questionnaire was taken by 61 participants. The data was collected from April 1st, 2023, to April 29th, 2023. The participants were over the age of 18 years and had a background in software security. Of the 61 participants, 44 were male and 17 were female. Question 1 asked the participants to select their age range. The possible age range choices were under 18, 18-24, 25-34, 35-44, 45-54, 55-64, and over 65. Of the participants, 0% (0 participants) were under 18, 8.2% (5 participants) were in the 18-24 range, 16.4% (10 participants) were in the 25-34 range, 21.3% (13 participants) were in the 35-44 range, 26.2% (16 participants) were in the 45-54 range, 16.4% (10 participants) were between the age of 55 and 64, and 11.5% (7 participants) were over the age of 65. The demographic factors of the participants who completed this survey can be found in Figure 1. The job status (academia, industry, both, and retired) of the participants who completed this survey can be found in Figure 2.

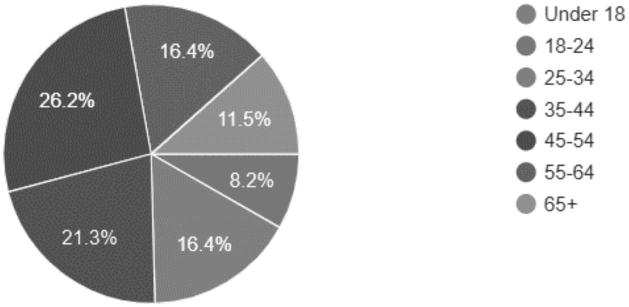


Figure 1: Demographics of the participants

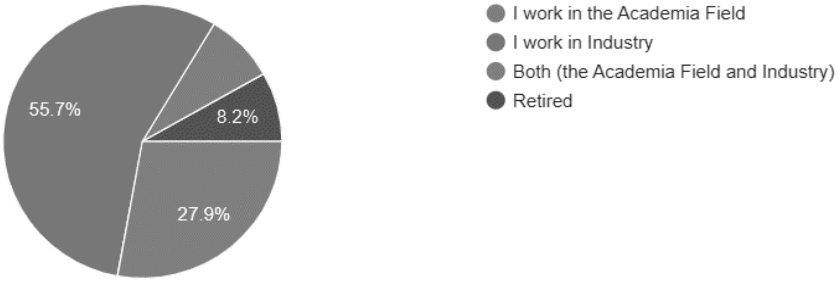


Figure 2: Job status of the participants

Findings

In research question 4 (RQ4), 83.7% of the participants selected “Strongly Agree” or “Agree” regarding their opinion that sending encrypted messages using a cryptographic software tool interface would be easy to use and understand. In research question 5 (RQ5), 47.5% of the participants selected “Strongly Disagree” or “Disagree” regarding their opinion that sending encrypted messages using a cryptographic software tool interface would NOT be easy to use and understand. In research question 6 (RQ6), 72.1% of the participants selected “Strongly Agree” or “Agree” regarding their opinion that all data and communication in every application should be encrypted. In research, question 7 (RQ7), 52.5% of the participants selected “Strongly Agree” or “Agree” regarding their opinion that sending encrypted messages using the command-line interface would be easy to use and understand.

In research question 8 (RQ8), 55.7% of the participants selected “Strongly Agree” or “Agree” regarding their opinion that sending encrypted messages using the command-line interface would NOT be easy to use and understand. In research, question 9 (RQ9), 85.3% of the participants selected “Strongly Agree” or “Agree” regarding their opinion that all data regarding private information should be confidential and should be a top priority when developing an application. In research, question 10 (RQ10), 73.8% of the participants selected “Strongly Agree” or “Agree” regarding their opinion that all private information should be stored on a server that is stored in the country where I reside.

It is important to observe that over 70% of the participants selected “Strongly Agree” or “Agree” for research questions 6 (RQ6) and 10 (RQ10). Another important observation was that over 80% of the participants selected “Strongly Agree” or “Agree” for research questions 5 (RQ5) and 9 (RQ9). Table 1 shows the breakdown of the survey results.

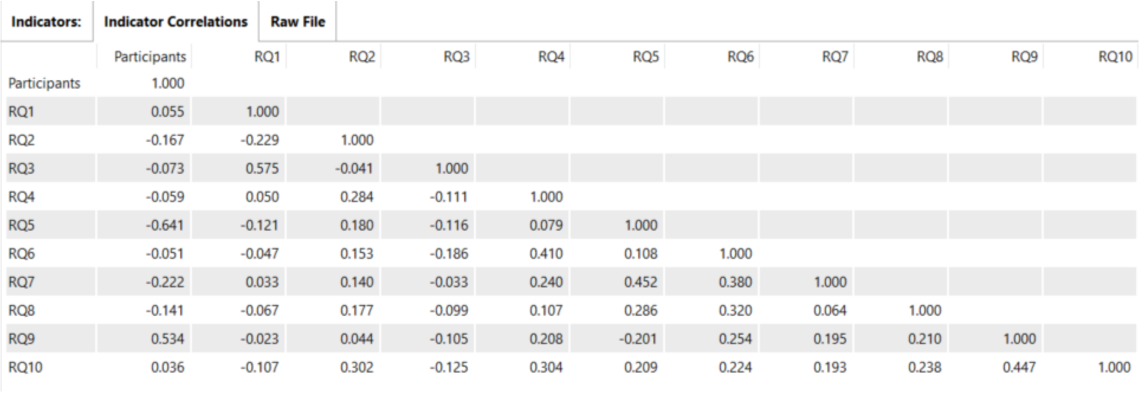
Table 1: Survey Results.

RQ	SA	A	NAND	D	SD
RQ_4	14	37	5	4	1
RQ_5	11	14	7	26	3
RQ_6	15	29	3	13	1
RQ_7	9	23	8	15	6
RQ_8	11	23	8	15	4
RQ_9	30	22	0	6	3
RQ_{10}	25	20	5	10	1

Data Synthesis for Research Question

After analyzing the data, we looked at the path coefficient and the associated p-values. The path coefficient weights indicated the following hypothesis (H1-H5): path coefficient weights < .001 indicates a small effect while path coefficient weight < 0.30 indicates a medium effect and path coefficient weight < 0.50 indicates a large effect. Generally, the path coefficients should be less than 1.000.

This study has provided information from students and SMEs working in the IT field about their usage of a cryptographic software tool for data encryption and communication within their line of work. The coefficients of the constructs have shown a significant correlation between the usage of a cryptographic tool to send encrypted data communications and the understanding of its importance among respondents. Therefore, all five hypotheses are supported. Figure 3 shows the indicator correlations between construct variables among respondents.



	Participants	RQ1	RQ2	RQ3	RQ4	RQ5	RQ6	RQ7	RQ8	RQ9	RQ10
Participants	1.000										
RQ1	0.055	1.000									
RQ2	-0.167	-0.229	1.000								
RQ3	-0.073	0.575	-0.041	1.000							
RQ4	-0.059	0.050	0.284	-0.111	1.000						
RQ5	-0.641	-0.121	0.180	-0.116	0.079	1.000					
RQ6	-0.051	-0.047	0.153	-0.186	0.410	0.108	1.000				
RQ7	-0.222	0.033	0.140	-0.033	0.240	0.452	0.380	1.000			
RQ8	-0.141	-0.067	0.177	-0.099	0.107	0.286	0.320	0.064	1.000		
RQ9	0.534	-0.023	0.044	-0.105	0.208	-0.201	0.254	0.195	0.210	1.000	
RQ10	0.036	-0.107	0.302	-0.125	0.304	0.209	0.224	0.193	0.238	0.447	1.000

Figure 3: Indicator correlations

Summary

The usage of a cryptographic tool used for data encryption and its effect on data privacy was discussed and reviewed to answer research question RQ1: Is it important to encrypt messages, information, and communication using cryptographic software? and RQ2: What is the level of difficulty in using a cryptographic software tool to send encrypted messages? These two research questions resulted in the first hypothesis (H1). This hypothesis was significant and had high path coefficient values. It was established that sending an encrypted message using modern operating systems can provide data encryption and have a positive effect on usage. The five other hypotheses (H1, H2, H3, H4, H5), were designed to measure the perception of usage of a cryptographic tool and data encryption for communication among college students and IT professionals. These five hypotheses were significant and had a high path coefficient which supported the theory and hypotheses. These hypotheses are supported by the strong correlation between variables indicator R4-R10. Figure 4 displays the path model and coefficients that was created using SmartPLS3.

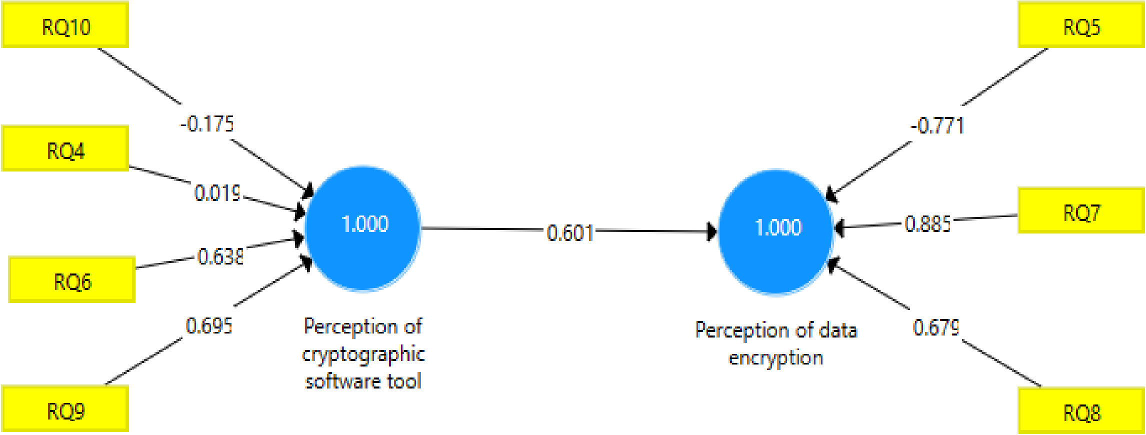


Figure 4: Path model using SmartPLS3

The analysis shows that hypothesis 1 is strongly supported with a significant (p<1.000) path coefficient weight of 0.60, which indicates a significant effect regarding the perception of the usage of the cryptographic tool and how it affects the perception of data encryption and communication among college students and IT professionals, and SMEs. This finding validates our preconceived notion of a positive correlation effect between a cryptographic tool and data encryption. Hypothesis 2: "The participants will perceive that sending encrypted messages using a cryptographic software tool will find that it is less intuitive than the command-line interface" has a positive correlation with the usage of a cryptographic tool and data encryption. The analysis shows that hypothesis 2 is strongly supported with a significant ($p<1.000$) path coefficient weight of 0.63, which indicates a significant effect among participants. Prior studies and literature agree that using a cryptographic tool for data encryption will increase the user's confidence in data communication.

This study suggests that cryptography and data encryption will increase end users' level of data security confidence when sending and receiving sensitive data. Hypothesis 3: "The participants will perceive that encrypting messages and data is necessary" The analysis shows that hypothesis 3 is strongly supported with a significant ($p<1.000$) path coefficient weight of 0.69, which indicates a significant effect. This validates the finding of this study regarding the preconceived notion of a positive correlation between a cryptographic tool and data encryption when sending and receiving data.

This study suggests that stronger cryptography and data encryption will increase users' confidence when sending and receiving sensitive data using different operating systems and network communication. Hypothesis 4 and Hypothesis 5: "The participants will perceive that all data and communications in every application should be encrypted" The analysis shows that hypothesis 4 is strongly supported with a significant ($p<1.000$) path coefficient weight of -0.17 which indicates a minor effect. This validates our preconceived notion of a positive correlation between the usage of a cryptographic tool for data encryption and communication. This study suggests that a stronger cryptographic tool and data encryption will increase users' perception of data security when sending and receiving sensitive data.

Reliability

To evaluate the reliability of the measurements, this study used the PLS calculations to ensure the quality of the theoretical framework by removing measurement errors such as poorly formulated questions. Also, reliability was established by using three different subject matter experts (SMEs) to generate the survey questions. A SME is an individual who is a specialist in their field, with degrees and years of experience in a particular topic (Mattoon, 2005). The SMEs made the determination of which questions should be on the questionnaire based on their knowledge and experience. The candidates to be SMEs in this research were recruited through a list of individuals in the software security field. The candidates were determined based on their experience working within their chosen field, collaboration techniques, and soft skills.

Discussion

This study identified factors that were reported in previous studies and thus confirms the measurements of usage of the cryptographic tool for data encryption and communication. Based on the literature review, the relationship between the usage of a cryptographic tool for data encryption and communication has not been empirically tested in the past. In addition, to the best of the researcher's knowledge, there are not many similar studies that mainly focus on IT professionals. The results displayed that 26 of the respondents answered that they disagreed with R5 "I believe that sending encrypted messages using a cryptographic Software tool would be easy to use and understand."

The findings indicated that a cryptographic technology tool can lead to reliance and dependence on the software than a solid data security judgment among respondents. A possible explanation for this finding could be because these respondents may not be working with data encryption technologies in their daily work. Some respondents may not have fully understood the implication of the question. The lack of terminology knowledge could also be a factor for why some respondents do not see any potential for a cryptographic tool and data encryption in their daily work.

Limitations

The implementation of this study was not without certain limitations. The study is limited by the fact that it only focused on the measurement of specific variables. A limitation exists regarding the personal views of the participants using a specific software tool to send encrypted messages. It is unknown whether certain answers to the questions were biased, based upon the participant's previous experience in the software security field, therefore altering the acceptance of one or more hypotheses. Another limitation is the lack of knowledge regarding the participants' experience in the software security field and the size of the data sample. Further investigation is needed to establish if the same results could be duplicated through a larger data sample and applied across a broader context.

A future study could include the expansion of the sample size, which would be warranted to eliminate as much bias in the responses as possible. The need to balance the study by age group may also be warranted to remove any significant influence on the survey. Lastly, a second survey could be conducted on the same participants after they have learned more information related to software security, asking the same questions to determine if their opinions had changed since the participants' previous responses.

Conclusion

This study aims to highlight the impact that respondents believed heavily that all data regarding private information should be confidential, should be a top priority when developing an application, and all private information should be stored on a server that is stored in the country where they reside. This mindset runs in parallel with General Data Protection Regulation (GDPR) European Union practices. The General Data Protection Regulation (GDPR) is a regulation in EU law on data protection and privacy for all individuals within the European Union (EU) and the European Economic Area (EEA). It also adopts the transfer of personal data outside the EU and EEA areas. GDPR seeks to undertake control of individuals over their data and to streamline the regulatory environment for international business by combining the regulation within the EU (Tankard, 2016).

GDPR brings into account the effect of "data sovereignty, data privacy, data localization, and data residency are four interconnected concepts that play a significant role in the modern digital landscape. As the global flow of information rapidly expands, businesses, governments, and individuals must navigate these complex issues to ensure compliance with regulations and protect their privacy and intellectual property" (Freestone, 2022). This is what US companies should keep in mind when dealing with data.

The overall results of this study supported the fact that respondents agreed that all data and communication in every application should be encrypted in an easy manner. This study demonstrated that a cryptographic technology tool can lead to reliance and dependence on a software tool rather than a solid data security judgment among respondents.

This study's results found that a stronger cryptographic tool and data encryption will increase users' perception of data security when sending and receiving sensitive data. In conclusion, the results show that stronger cryptography and data encryption will increase users' confidence when sending and receiving sensitive data using different operating systems and network communication.

References

Anugurala, A., & Chopra, A. (2016). Securing and preventing man in middle attack in grid using open pretty good privacy (PGP). In *2016 Fourth International Conference on Parallel, Distributed and Grid Computing (PDGC)* (pp. 517-521). IEEE.

Bajaj, P., Arora, R., Khurana, M., & Mahajan, S. (2022). Cloud security: The future of data storage. In *Cyber Security and Digital Forensics: Proceedings of ICCSDF 2021* (pp. 87-98). Springer Singapore.

Dechand, S., Naiakshina, A., Danilova, A., & Smith, M. (2019). In encryption we don't trust: The effect of end-to-end encryption to the masses on user perception. In *2019 IEEE European Symposium on Security and Privacy (EuroS&P)* (pp. 401-415). IEEE.

Dizon, M. A. C. (2023). The value of trust in encryption: Impact and implications on technology law and policy. *IEEE Transactions on Technology and Society*.

Freestone, T. (2022, November 15). *Data Sovereignty and GDPR [Understanding Data Security].*

Kiteworks. https://www.kiteworks.com/gdpr-compliance/data-sovereignty-gdpr/

Johnson, J., Houghton, R., Hilton, T., & Cheah, K. F. (2018). Secure data communication via lingual transformation.

Kang, H., & Deng, J. (2023). A cross encryption scheme for data security storage in cloud computing environment. *International Journal of Internet Protocol Technology*, *16*(1), 1-10.

Kumar, D., Kashyap, D., Mishra, K. K., & Misra, A. K. (2010). Security Vs cost: An issue of multi-objective optimization for choosing PGP algorithms. In *2010 International conference on computer and communication technology (ICCCT)* (pp. 532-535). IEEE.

Papoutsakis, M., Fysarakis, K., Spanoudakis, G., Ioannidis, S., & Koloutsou, K. (2021). Towards a collection of security and privacy patterns. *Applied Sciences*, *11*(4), 1396.

Reuter, A., Abdelmaksoud, A., Boudaoud, K., & Winckler, M. (2021). Usability of end-to-end encryption in e-mail communication. *Frontiers in big Data*, *4*, 568284.

Shen, Y. (2021). End-to-end encrypted messaging based on PGP with forward secrecy. In *Journal of Physics: Conference Series* (Vol. 1873, No. 1, p. 012031). IOP Publishing.

Tankard, C. (2016). What the GDPR means for businesses. *In Journal of Network Security* (Vol. 2016, No. 6, p. 5-8).