Faifr, Adam; Januška, Martin

Article

Factors determining the extent of GDPR implementation within organizations: Empirical evidence from Czech Republic

Journal of Business Economics and Management (JBEM)

Provided in Cooperation with:
Vilnius Gediminas Technical University

Suggested Citation: Faifr, Adam; Januška, Martin (2021) : Factors determining the extent of GDPR implementation within organizations: Empirical evidence from Czech Republic, Journal of Business Economics and Management (JBEM), ISSN 2029-4433, Vilnius Gediminas Technical University, Vilnius, Vol. 22, Iss. 5, pp. 1124-1141, https://doi.org/10.3846/jbem.2021.15095

This Version is available at:
https://hdl.handle.net/10419/317515

VILNIUS TECH
Vilnius Gediminas Technical University

Journal of Business Economics and Management
ISSN 1611-1699 / eISSN 2029-4433
2021 Volume 22 Issue 5: 1124–1141
https://doi.org/10.3846/jbem.2021.15095

FACTORS DETERMINING THE EXTENT OF GDPR IMPLEMENTATION WITHIN ORGANIZATIONS: EMPIRICAL EVIDENCE FROM CZECH REPUBLIC

Adam FAIFR*, Martin JANUŠKA

Department of Business Administration and Management, Faculty of Economics, University of West Bohemia, Univerzitní 22, 306 14 Pilsen, Czech Republic

Received 02 June 2020; accepted 01 April 2021

Abstract. In this paper, the key factors that affect the extent of GDPR implementation in enterprises are analysed. Since 2018, all organizations operating in the European Union or processing personal data of EU citizens have had to incorporate a new regulation in their work. After three years of experience, possible key factors that significantly affect the cost of the entire project have been theoretically identified. However, a research gap remains whether the factors thus defined actually have a real impact on the implementation within organizations. Therefore, this study focuses on an empirical investigation of those characteristics using quantitative approach combining Chi-squared tests and the Classification and Regression Tree method. Based on a survey of organizations in the Czech Republic, this paper outlines that the size of the organization, the typology of personal data processed and the way GDPR is implemented determine the scope of the implementation project within organizations. On the other hand, there is no clear evidence that there is significant role in whether it is a public or private organization.

Keywords: General Data Protection Regulation, GDPR, SMEs, implementation, organizations, compliance, public administration.

JEL Classification: M10, C89, L30, L50.

Introduction

In May 2016, the EU adopted a regulation that gives all citizens of the European Union greater protection of their personal data. The regulation known as the General Data Protection Regulation (GDPR) came into force in May 2018.

Although the GDPR legitimately increases the right of citizens to control their personal data, obviously, these claims have an impact on the operations of organizations that are facing a challenge of how to correctly implement the regulation in its enterprise or otherwise face the threat of fines (European Parliament, & Council of the European Union, 2016). For

*Corresponding author. E-mail: *faifr.adam@gmail.com*

many organizations, the challenge is to transform hundreds of pages of legal text into practical changes (Garber, 2018).

Based on literature research, it can be concluded that due to the recency of this topic since 2016, there has been a significant increase in interest not only in the academic but also in the commercial field. However, with regards to the research of the real impacts on organizations, the majority of publications are opinion pieces which discuss the possible consequences.

But there is little empirical evidence in this scientific area (Yuan & Li, 2019).

Although, an universally valid framework (regardless of the number of implementation steps taken) has been identified for any organization to achieve compliance (Almeida Teixeira et al., 2019; Garber, 2018; Tamburri, 2020), each GDPR project takes place in the context of given organization. Therefore, the scope, time frame and cost intensity of the implementation projects are always given by the contextual factors – specifically by characteristics of the selected organization as well as the specifics and circumstances of the implementation itself.

The main aim of this paper is to empirically investigate the real impact of theoretically defined contextual factors specifically on implementation costs. The accompanying part of the investigation is also the comparison of the significance of each examined factors and creation of effective hierarchy between them. Due to the fact that the Regulation itself allows Member States to adapt certain articles to their own legislative context, only organizations operating in the Czech Republic are the subject of this examination to prevent potential distortion.

From academic point of view, this output should complement the ongoing and (so far) mostly theoretically oriented research into the impact of the GDPR on organizations by filling the current research gap in terms of empirical verification. At the same time, the output should enable the business units to identify their own benchmarks regarding to GDPR implementation and estimate the complexity of the project.

This article is divided into chapters as follows: Section 1 summarizes the findings to date in relation to the impact of GDPR on organizations and summarizes the variables identified so far. Section 2 presents the methodological approach by which the subject will be analysed. Section 3 is devoted to the presentation of empirical results and its analysis, while in the final section all known findings are summarized, the limitations of this research are mentioned, including proposals for further research.

1. Literature review and propositions

As implied by the nature of the Regulation taken, and even on the basis of some investigations and opinions already published (Khan, 2018; Hoofnagle et al., 2019; Almeida Teixeira et al., 2019), the implementation of GDPR is not supposed to be a short-term one-off activity, but a complex activity that fully becomes the part of any organization that in any way processes personal data – either about customers or about their own employees (Perry, 2019). Garber (2018) also states: "*If a business does not have a sufficient overview of its data, this process will disrupt the IT department, compliance team and the business itself.*"

Given the wording of the Regulation, which emphasizes the responsibility of processors in the management of personal data, implementation itself requires a comprehensive approach that appears to affect both procedures relating to data processing as well as procedures

without such a relation (Hoofnagle et al., 2019). In order to ensure sustainability of compliance with the Regulation, all processes need to be reviewed and an appropriate structure needs to be established to meet all the requirements of the Regulation (Tikkinen-Piri et al., 2018).

With a nearly three years of experience after the commencement of GDPR, it is clear that, despite the threat of large sanctions, not all businesses have been able to implement the Regulation as a whole on time (Garber, 2018). The main reason for the failure of individual companies to implement the Regulation is the need for a comprehensive solution. Legislation itself plays a part in this, because there is no statement on a how particular organization should achieve the desired state. No prescription of technology is provided (Tankard, 2016). Also, as Yuan and Li (2019) state, for many organizations, changes have required and still require costly solutions that have had a negative impact on their economic performance in the short term. Although financial costs may appear to be the alpha and omega of the success of the entire implementation, there is also a need for time, staff training as well as changes in strategic planning concepts (Tikkinen-Piri et al., 2018).

The implementation of GDPR cannot be seen only as a process that aggravates and complicates the running of a business. Despite the challenge for each organization to change its own data management reorganization (Hofman et al., 2019) and changes that have a temporary negative impact on business results, the right approach to the implementation can bring several positive effects in the long run. By mapping its own data, the organization gets an idea of the overall information flow. Its optimization can lead to lower operational costs (Beckett, 2017; Perry, 2019), more credible data analysis (Garber, 2018; Almeida Teixeira et al., 2019) and, last but not least, increased customer confidence and corporate competitiveness (Beckett, 2017; Almeida Teixeira et al., 2019; Datoo, 2018).

1.1. Implementation framework

As mentioned above, the Personal Data Protection Regulation does not precisely define the technological means and procedures that should lead to its successful incorporation (Tankard, 2016; Tamburri, 2020). Based on a broad systematic literature review (Almeida Teixeira et al., 2019) published in the period after the commencement of GDPR, individual articles propose a three- or four-step implementation procedure. Regardless of the number of steps, the individual subframes coincide in the procedure.

The first and fundamental step is a comprehensive understanding of the organization in terms of structure and continuity of individual data (Almeida Teixeira et al., 2019). Every processing takes place within a certain system framework and therefore it is necessary to take into account the entire data management of the company, all the systems used, the way they are used and the analysis of responsibility for data collection and analysis. GDPR does not limit its scope to the organization itself, but considers also the issue of potential cooperation with third parties (Hoofnagle et al., 2019; Starčevič et al., 2018; Udroiu et al., 2018).

The enterprise architecture mapping process is used for gap analysis to identify bottlenecks throughout the system that prevent a company from being compliant. By its nature, this section does not apply only to personal data, but to all processed data. A positive side effect may be the identification of duplicities (Udroiu et al., 2018), which goes hand in hand

with the principle of regulation based on minimizing the processed data (European Parliament, & Council of the European Union, 2016). The regulation itself also considers DPIA, which can assess current risks. Although this step is required for cases where there is a high risk for the data subject, for example Kindt (2018) recommends to perform this assessment in any case.

The implementation is a multidisciplinary issue where it is necessary to take into account the intended purpose and goal of the project (project management), minimize the negative consequences and prioritize the individual steps with respect to future risks of data leakage and organization insufficiency (risk management) or the creation of an organizational climate that will be able to implement the changes as well as follow the principles even after the implementation itself (change management) (Perry, 2019).

1.2. SMEs and large companies

Further analysis focuses on comparing the whole implementation process of enterprises according to their size. Many studies carried out before the Regulation came into force identified many factors that prevent organizations from implementing them (Almeida Teixeira et al., 2019). The size of the organization seems to be a key factor, as it appears even after its commencement. While large companies have seen the implementation of GDPR as a reasonable and viable solution from the outset (Sirur et al., 2018), for smaller businesses the change itself is associated with larger organizational changes that go beyond the organization's existing knowledge base (Lindgren, 2018; Perry, 2019).

Based not only on studies published after its commencement in the UK (e.g. Garber, 2018), the general preparedness of the organization is also related to the size of the organization. A significant proportion of smaller companies did not manage to make the appropriate changes, while for larger companies the problem was significantly smaller (Perry, 2019).

However, the factor of the size of the organization has to be more precisely understood with regard to the related general maturity of individual organization in managing their own data – knowledge of data management and privacy knowledge. Also, the amount of free funds to implement has to be considered. As mentioned above, the Regulation itself provides only qualitative statements. Consequently, while larger companies have been able to translate flexible interpretations into specific steps, the large number of small companies that were not able to implement GDPR on time suggests a potential struggle with understanding the individual requirements and consequential implementation (Garber, 2018; Perry, 2019).

Implementation within SMEs is associated with major process changes (Lindgren, 2018) and these organizations at the same time are more vulnerable to fraudulent practices of external services, which above all pose a higher risk with regards to possible leaks and protection of data subjects' rights (Longras et al., 2018; Sirur et al., 2018).

In terms of their size, there are two types of organizations. Those that are able to meet all demands, they create an environment for further development and strengthen their confidence, and those organizations that are groping with the sub-regulations will continue to fumble and lose potential for growth and possibly risk the confidence of their current customers, and also future ones (Martin et al., 2020; Datoo, 2018; Yuan & Li, 2019).

Regarding to Bleier et al. (2020), this aspect may be considered by policy makers since the well-intended regulations (such as GDPR) might have undesired downstream effects. Gal and Aviv (2020) also mention similar unintended and so far unrecognized effects on competition, efficiency or innovation.

To a certain degree, the inabilities resulting from the lower maturity of the organization can be mitigated by the help of public authorities, but this kind of assistance was considered insufficient before the Regulation came into effect, according to one qualitative research (Sirur et al., 2018).

Proposition 1: Large companies have invested more money in the implementation of GDPR than SMEs

As mentioned above, two basic groups of organizations can be distinguished, where the main dividing characteristic is the size of the organization in terms of the number of employees. With increasing size of the company, the data and general complexity are possibly to increase as well as the scope of implementation project.

On the other hand, due to the lack of empirical evidence supporting the defined proposition, with regards to possible moderating effects, this paper considers the already examined similarity of GDPR requirements with the requirements of ISO 27K and other related standards (Mesquida & Mas, 2015; Diamantopoulou et al., 2020). As Mesquida and Mas (2015) outline in case of precedent ISO 27K certification, acquisition of additional standard is assumed to be less time- and cost- consuming given by previous acquirement of knowledge and necessary capabilities. In the field of GDPR implementation, Diamantopoulou et al. (2019) state that some companies certified by ISO 27K standards used joint synergy effects to speed up the implementation of GDPR. In some other cases, implementation of GDPR was facilitated by previously implemented ISO 27000 (Longras et al., 2018).

Regarding to the defined proposition, it is then necessary to conclude that the above-mentioned and knowledge-related synergies in general will have an effect especially among larger companies, which are more often to be certified (Everett, 2011). This assumption is also supported by the need for major process changes among SMEs (Lindgren, 2018).

Notwithstanding the above, research further points to the fact that project complexity is not necessarily correlated with rising project costs (Nguyen et al., 2019). In other words, implementing GDPR in a more complex environment would not necessarily mean more expensive implementation.

Nevertheless, since it is not possible to satisfactorily assess whether the mentioned moderation effect works absolutely or only relatively (with regards to the annual sales or annual organizational costs), default proposition assumes implementation in a larger company environment to be associated with a greater time, organizational and financial investments than is the case in small and medium-sized companies.

Proposition 2: There are differences in the implementation of GDPR internally and in cooperation with an external entity

The internal implementation of GDPR may be associated with the need to acquire new knowledge and skills (Perry, 2019; Garber, 2018), which may be reflected in the overall cost

of implementation. In addition, a substantial part of companies consists of micro and small enterprises, which are considered less capable (Sirur et al., 2018) and which will rather address an external entity for implementation. However, it is not yet clarified which of the two strategies mentioned is more cost effective for implementation, even with regards to the size and needs of the organization.

Although in view of the above mentioned, differences between the two strategies can be expected, but specific differences and regularities between them have not yet been investigated.

1.3. Sensitive and non-sensitive data

Particular attention is paid to the differentiation of personal data in the Regulation, where the so-called special categories of personal data are newly explicitly defined as follows: *"personal data membership, and processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data relating to health or data relating to a natural person's sex life or sexual orientation shall be prohibited"* (European Parliament, & Council of the European Union, 2016).

As a result, these types of personal data are gaining more attention. The protection of data subjects when processing such data is legally enhanced (Yuan & Li, 2019; Larrucea et al., 2020). However, this also has practical implications for implementation if the company processes such data. Sensitive data, by its nature, increases the risk of leakage or inadequate processing (Prakash & Singaravel, 2015). With increasing risk, attention should then be directed in such a way and to the extent that the risk to the data subject is minimal (Hoofnagle et al., 2019).

Accordingly, in the case of the processing of data that fall into a special category, the Regulation requires the compilation of a Data protection impact assessment (European Parliament, & Council of the European Union, 2016; Quinn & Quinn, 2018). This will undoubtedly affect the time and overall implementation costs. Healthcare facilities that process patient health information may be an example. Research and studies of this environment point to costly investments in preparation to meet all the requirements (Quinn & Quinn, 2018; Yuan & Li, 2019).

However, this fact can be also theoretically implied on all other entities that in some way process data within a special category. They may also be data processing entities for children (Starčevič et al., 2018; Tikkinen-Pirri et al., 2018). Any processing in this respect must be duly justified and treated with special technical measures (Starčevič et al., 2018). In a special case of risky processing of personal data, information can then be obtained by IoT devices operating, e.g., with biometric data, which also fall into the special category (Kindt, 2018; Kounoudes & Kapitsaki, 2020).

Proposition 3: Organizations that process personal data in a special category invest more in implementation than other organizations

The processing of sensitive personal data is associated with a higher risk (European Parliament, & Council of the European Union, 2016). As a result, the whole process of GDPR

implementation is more complex and at the same time more costly than for companies that do not keep personal data in this category.

Proposition 4: Public institutions invest more in the implementation of GDPR than private entities

The management of personal data and, in many cases, sensitive data is a common agenda for public institutions, while private entities whose main task is not personal data management do not normally process these data.

2. Data and methods

Given the nature of this research and the definition of the research question, this research has an exploratory character where quantitative research methods are used for evaluation. As mentioned above, the aim of this research is to empirically investigate the factors affecting the overall cost of the GDPR implementation project in organizations. These factors are defined in more detail on the basis of a previous literature review. In the academic field, however, the strength and hierarchy of relationships of individual characteristics have not yet been measured.

In the selection of suitable variables, however, in the initial phase, the linkage of individual independent variables to the scope of GDPR implementation will be examined. Due to the categorical nature of the input data, this will be done on the basis of defined variables with respect to propositions by individual chi-square tests based on which the propositions will be confirmed or rejected. The chi-square test of data independence is based on a comparison of obtained and expected frequencies of occurrence of a given variable. This type of test is used to determine whether the relationship between defined independent variables and dependent variable is significant (Huber-Carol et al., 2002; Sirkin, 2006).

Subsequently, the data mining model will be used to create an efficient hierarchy, namely the Classification and Regression Tree (C&RT) showing the best prediction with an accuracy of 80.81% among data mining models (Park et al., 2013). Classification and regression trees (CART) are a non-parametric decision-tree learning technique that produces either classification or regression trees, depending on whether the dependent variable is categorical or numeric, respectively (Strickland, 2016).

2.1. Sample and data collection

Empirical analysis is based on the results of survey conducted within closed population of Czech organizations (Maňourova, 2019). Conduct survey on closed population is, for example, recommended by Sue and Ritter (2007). In this regard, only two basic conditions for targeted group have been defined – the organization processes the personal data of EU citizens and the organization is a public or private unit operating in the Czech Republic. Since the key characteristics of sample are known, multistage (also known as "clustering") procedure was selected (Creswell, 2013).

After identifying the target groups of respondents, according to Sue and Ritter (2007), the next stage is to generate a list of the population members (sampling frame). Compared to previous legislation (The office for personal data protection, 2018), there is no database associating all entities processing personal data. In this context, publicly available databases were used to create a sufficiently comprehensive list. In case of public entities, these were databases of relevant Ministries (Ministry of the Interior, Ministry of Education, Ministry of Health). In case of private entities, the selection was made on the basis of publicly available catalogues of private organizations, where it was subsequently verified that it is an active entity corresponding to defined criteria. Overall sampling frame comprised 1500 contacts. To ensure the same probability of being in the survey, the sample included equally both public and private entities.

Subsequently, individual representatives with appropriate insight into the implementation of GDPR in the given entity, were addressed. In this context, the relevance of the contact addresses regarding to two elements of eligibility (Sue & Ritter, 2007) was also considered. Further, prior to submitting the questionnaire electronically, respondents were informed of the content and objectives of the survey via email.

The research is based on results obtained from a total of 223 Czech organizations. After exclusion of apparently ineligible results, the response rate is 14.87%. 175 respondents held the position of executive in the organization (Mayor, Principal, CEO…), the remaining respondents were employees directly related to the implementation of GDPR in the company (Data protection officer, senior management, lawyer, secretary ...).

The profile of organizations according to the relevant criteria is provided in Table 1.

Table 1. Information about organizations reviewed

Criteria	Characteristic	Number
Size of company	Micro-enterprise	115
	Small enterprise	60
	Medium enterprise	34
	Large enterprise	14
Legal entity	Public entity	163
	Private entity	60
Type of organization	Self-government, self-governing office and unit	130
	Educational institution	32
	Manufacturing company	30
	Service company	20
	Trade and brokering company	10
	Hospital	1

It follows from the above that majority of respondents consisted of organizations whose size corresponds to a micro-enterprise These were mainly public institutions, mainly local authorities. The authors of this article see the higher rate of return of the questionnaire from public entities generally in a more formalized and transparent approach to the implementation of GDPR than in the case of private organizations.

To prevent possible bias in results due to the characterisctics of the collected sample, partial chi-square tests ale also performed between the selected independent variables. The results of those tests are presented together with the main propositions examined.

2.2. Dependent variable

The overall impact of GDPR implementation on organizations can be measured from several different angles – implementation time, implementation costs, changes in business performance, business size, etc. As this article looks for key characteristics that determine the entire implementation process, the direct calculated cost of the implementation was chosen as the key metric.

A total of 5 different cost ranges were chosen for the research and were ordinarily divided in the analysis. A higher score means higher implementation costs. The options included: 1 – "less than 10 000 CZK", 2 – "10 000–50 000 CZK", 3 – "50 000–100 000 CZK", 4 – "100 000–200 000 CZK" and 5 – "more than 200,000 CZK". Within the interval, the respondent chose only one option. The categories used in this way are based on a larger survey conducted in the Czech Republic. As the Czech Chamber of Commerce stated – according to its investigations – a statutory representative of Czech entrepreneurs: "*Most companies will pay up to CZK 50,000 for GDPR. Each of more than one-fifth of large companies (i.e., with more than 250 employees) spent over 500 000 CZK on these measures. A fifth of the companies with up to 10 employees also said they had no spending on GDPR because they had not prepared for it in any particular way. These companies either believe that the new regulation will not affect them or rely on avoiding the consequences*" (Czech Chamber of Commerce, 2018).

2.3. Independent variables

The analysis follows the key characteristics that determine the extent of the entire GDPR implementation. In the Czech environment, and also on the basis of a literature review, there is no research that defines the determining characteristics. Literary research carried out in the previous section however states the expected variables. The identified variables include in particular:

Size of the company

While the degree of maturity can be measured more complicatedly, in this connection the size of the enterprise gives a clear view. In the UK, for example, studies have revealed problems for smaller companies in implementing GDPR at all (Garber, 2018). Within the data collection, the categories of enterprises were defined according to the methodology outlined by the Council of the European Union (2013): "Micro-enterprise – up to 10 employees"; "Small business – up to 50 employees"; "Medium Enterprise – Fewer than 250 Employees" and "Large Enterprise – More than 250 Employees". The ascending ordinal approach was used ranging from 1 to 4.

Implementation approach

There are several approaches to how to implement GDPR. These include, inter alia, the factor of whether everything is done internally or, to a greater or lesser extent, in cooperation with an external entity. According to research, organizations with insufficient privacy knowledge may resort to external cooperation (Lindgren, 2018). Secondly, it may also be that larger companies are able to invest higher budgets in implementation (Sirur et al., 2018), which may be a combined form of strategy.

Two options were defined in the questionnaire: "Internal only" and "In cooperation with an external entity". This variable was defined as dichotomous, that is: "0 – internally" and "1 – external entity role". For replenishment, during the data collection, only 12 out of 224 respondents stated they would not receive any external offer for implementation. In half of the cases there were 2–4 offers for cooperation.

Sensitivity of personal data

The second dichotomous variable was whether companies, within their own scope, also process sensitive data about their clients. The options were defined as follows: "0 – the organization does not process any sensitive personal data" or "1 – the organization also processes sensitive personal data". The definition of this variable responds to the regulation itself, where the special new categories are defined either in the finding that processing sensitive data means a more demanding implementation process (Yuan & Li, 2019; McCall, 2018).

Public vs. private

The last factor is the influence of the legal entity of the organization, i.e., whether it is a private or public organization. It is for comparison marked as "O – Public entity" and "1 – Private entity".

As defined in Act No. 110/2019, § 61, Article 3 (Parliament of the Czech Republic, 2019): "*The Office shall refrain from imposing an administrative penalty also in the case of entities referred to in Article 83 (7) of Regulation (EU) 2016 / 679.*"

Thus, in the Czech legal environment there are differences in conditions between public and private entities, where public entities are not at risk of being imposed a fine in case of violation of the rules defined in the Regulation.

3. Results and discussion

In accordance with the presented methodology, the results of the tests are presented in this section. First, the results of tests of partial characteristics are presented and then their interconnections are outlined.

3.1. Characteristics determining the extent of implementation

The following Table 2 presents the results of partial tests. The first independent variable defined was the size of the organization and its relation to the cost of GDPR implementation.

Tests of independence revealed that there is a link between the size of the business and the expenditure that the company invests in implementation. In the case of small enterprises, more than half of the companies invested an amount of up to CZK 10,000 and in 80% the costs did not exceed CZK 50,000. In a detailed examination of micro-enterprises (organizations up to 10 employees) the shares were even higher in both cases. This finding supports the previous results of the Czech Chamber of Commerce (2018).

Significant differences in costs can be further seen in comparison with medium-sized and large enterprises. While in the case of medium-sized enterprises the costs are largely up to 100,000 CZK, in the case of large organizations most of the companies spent more than 200,000 Czech crowns in the process of launching.

Despite the considered moderating effects, the results do not refute the defined Proposition 1 that the cost of implementing GDPR increases with the size of the organization. However, the results are not entirely linear, when, for example among micro and small organizations, a significant difference between these two groups was not confirmed ($\alpha = .05$). At the same time, it is not possible to completely rule out moderating effects resulting, on the one hand from larger (and thus expected more expensive) organizational changes in small organizations (Lindgren, 2018; Perry, 2019) or on the other hand by proportionally cheaper implementation in larger organizations regarding to precedent implementation of related ISO standards (Diamantopoulou et al., 2019; Longras et al., 2018) or higher general capability considered (Garber, 2018; Perry, 2019). The moderating effect can still have an effect, but rather relatively.

Table 2. Test of independence results

Tested criterion	X^2	df	sample size	$\alpha = .05$	p-value
Size of company (default grouping)	90.87	12	223	21.03	0.000
Size of company (only Micro and Small enterprises)	4.28	4	175	9.49	0.369
Size of company (Micro and Small enterprises grouped together)	89.71	8	223	15.51	0.000
Sensitivity of data	15.6	4	223	9.49	0.003
Strategy of implementation	11.63	4	223	9.49	0.02
Type of organization	2.86	4	223	9.49	0.582

The second proposition reflects possible differences when an organization implemented GDPR internally or in cooperation with an external entity. The extent of this cooperation was not investigated. The test results show that both strategies offer differences. The null hypothesis in this case ($\alpha = 0.05$) is rejected. The p-value is 0.02. On closer inspection, more than 86% of internal implementations were budgeted for less than 50,000 CZK. For the second group of organizations, this interval moved to 100,000 CZK.

One possible explanation may be the fact that higher costs are primarily related to the size of the organization. To prevent a possible bias, an additional test was conducted to determine whether certain group of organizations were more likely to hire external collabora-

tors. However, the results obtained (X^2 (3, N = 223) = 4.58, p < 0.2) may not support this option. Neither the considered insufficient capabilities of smaller organizations leading to more frequent cooperation with external services (Lindgren, 2018), nor on the other hand the amount of available budget for implementation (Sirur et al., 2018) do not directly affect whether the organization will cooperate with an external entity. Subsequently, regarding to the defined proposition, it can be clarified that external cooperation is associated with the costs incurred regardless of the size of the organization.

The Regulation itself imposes additional obligations on companies processing personal data defined in special categories (European Parliament, & Council of the European Union, 2016; Quinn & Quinn, 2018). As shown in Table 2, additional obligations are analogous to higher costs. The conclusions presented by Yuan and Li (2019) can be confirmed not only in the public health sector but also in the public sector. However, even in the case of companies processing sensitive data, the fulfilment of obligations amounted to less than 50,000 CZK in approximately half of the cases. There is, however, a significant number of organizations whose costs exceeded the limit of 200,000 CZK.

The last variable was a comparison of implementation within public and private entities. There are different rules for these groups in Czech legislation. In general, both groups process personal data for other purposes. Table 2 shows no significant differences between these groups (p-value = 0.582).

Table 3. Comparison of public and private organizations

Type of organization	Total number of organizations	Share of organizations processing sensitive data	Share of cooperating organizations
Public organization	163	34.4%	83.4%
Private organization	60	6.7%	36.7%

As shown in Table 3, no significant differences were noted between the types of organizations, despite the fact that public organizations significantly more often process sensitive data and at the same time are more likely to cooperate with external subject during the implementation. Both above mentioned factors are also considered to be positively correlated with the amount of total costs (Table 2).

One of the ways to explain this contrast may be the fact that Czech public institutions fell under the scope of legislative requirements governing the issue of personal data management before GDPR was ratified (Nonnemann, 2011). This would mean that the process of adapting to GDPR might not have been that disruptive.

Another possible explanation is the form of cooperation with external entities, where due to the relative homogeneity of public organizations, the price offer of external entities is reduced. At the same time, in some cases, the organizations cooperated with the superior public entity in the implementation. However, none of these hypotheses have yet been examined.

All key findings from this chapter are summarized in the following Table 4.

Table 4. Findings summary

Proposition no.	Statement	Finding
1	Large companies have invested more money in the implementation of GDPR than SMEs.	At the normal level of significance ($\alpha = 0.05$), differences were proved.
2	There are differences in the implementation of GDPR internally and in cooperation with an external entity.	At the normal level of significance ($\alpha = 0.05$), differences were proved.
3	Organizations that process personal data in a special category invest more in implementation than other companies.	At the normal level of significance ($\alpha = 0.05$), differences were proved.
4	Public institutions invest more money in the implementation of GDPR than private entities.	No significant differences were found at the normal level of significance level ($\alpha = 0.05$).

3.2. Further analysis of interconnection of individual characteristics

Following the correlation between the individual characteristics of the organization and the scope of implementation activities of companies, the next part tries to find archetypal features and patterns having a key impact on the costs of meeting the requirements arising from the GDPR. It is assumed that the implementation itself is a multifactorial issue. In this context the decision tree method was used to identify variable clusters, in this case, groups of organizations with approximately the same costs invested in the implementation of GDPR. Based on the results, Figure 1 presents the decision tree, where a total of 7 non-terminal nodes (dividing characteristics) and 8 terminal nodes (clusters) were identified.

As it turns out, the size of an organization is a prerequisite for the overall scope of a GDPR implementation project. As outlined above, there are significant differences between micro and small organizations (ID = 3) and medium- and large-sized organizations (ID = 2). Given the relatively low robustness of the data obtained from medium-sized and large enterprises, these two groups do not break up further. In the case of groups of micro and small organizations, there is a further disintegration based on the implementation strategy, where it is possible to distinguish some organizations implementing internally (ID = 9) and in cooperation with external entities. For internal organizations, the implementation was associated with the minimum investment required, which is confirmed by the embedded histogram.

Another dividing variable is the type of personal data processed (ID = 8), broken down into organizations processing sensitive personal data and organizations processing standard data. The group of organizations processing sensitive data are primarily public institutions (local authorities, schools, etc.), while the remainder are private and public institutions with a lower degree of authority to the data management parties.

From the presented graph it can be argued that the primary determining factor is the size of the organization itself, which is related to the extent and complexity of the processed personal data and the size of the budget dedicated to the implementation itself. The secondary dividing factor is the implementation strategy and subsequently the question of

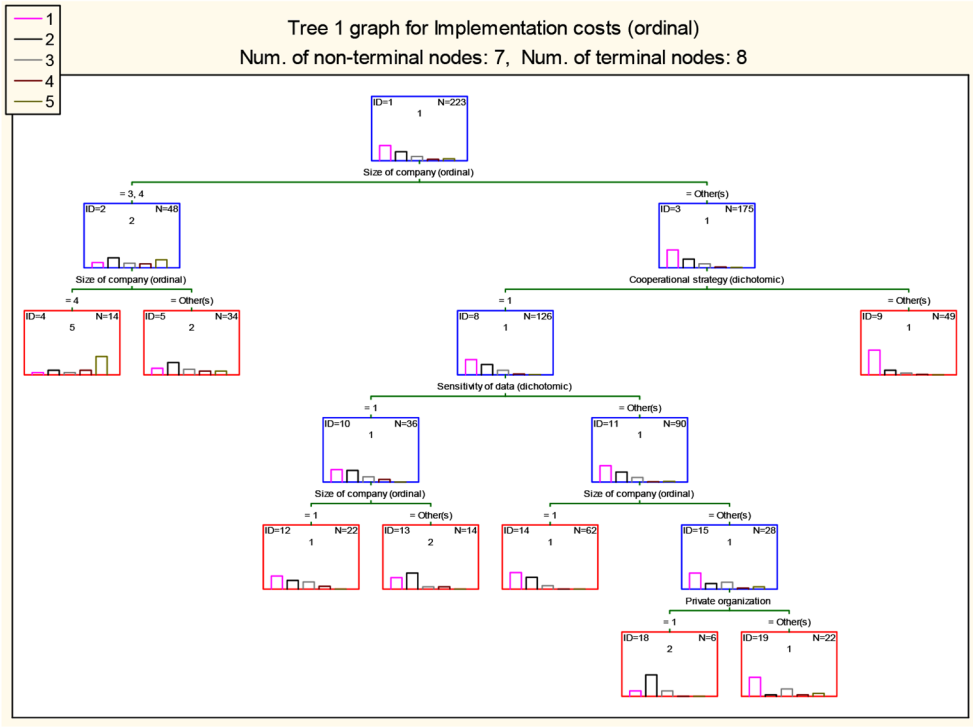

Figure 1. Correlation and regression tree (CART)

the sensitivity of personal data. This model verifies previous findings where the difference between public and private entities was not significant.

Conclusions

This article dealt with the implementation of GDPR in Czech organizations and aimed to identify the key factors that can determine the scope and cost intensity of the entire GDPR implementation.

As mentioned in the preparatory phase, each implementation of GDPR in an organization is a unique project and is based on a specific organizational structure, culture, different working and technological capital. The Regulation itself does not outline the procedures to be followed by organizations for its implementation. However, despite all the differences, common steps can be found to be taken, regardless of the organization. Although none of the sub-characteristics alone determines the exact scenario, by analysing their interconnections the individual segments of organizations can be identified.

Within this paper, key partitioning factors were identified on the basis of theoretical knowledge and their effects on implementation projects were subsequently examined. The extent of each GDPR implementation is determined by the size of the enterprise, the type of personal data processed, the implementation strategy that the organization incorporates.

The size of the organization has subsequent links to the quantitative and qualitative extent of the processed personal data. It can therefore be said with certainty that large organizations invest more money in the implementation of GDPR than medium and small enterprises. In the context of the research, it can also be stated that the size of the organization is the most important factor determining the implementation of GDPR. On the other hand, this relationship is clearly not directly proportional, as, for example, no significant differences have been demonstrated between micro and small enterprises. As the authors of this article outline, in this case some moderating variables may affect the cost and complexity of the entire implementation.

Another key criterion is also individual types of personal data. The Regulation attaches greater protection to so-called sensitive personal data. Higher protection in this case also means a technologically more advanced. The third examined factor is the method of implementation. This, as the research has shown, is more expensive when using an external subject than in the case of an internal solution. On the other hand, the quality of the implementation itself was not part of this research.

On the other hand, no differences were demonstrated between the groups of public and private institutions. However, despite fact that public institutions more often process sensitive personal data and more often cooperate with external entities during the implementation. Although the authors of this article provide possible explanations, attention to this specific phenomenon should be given in future research.

From the theoretical and academic point of view, this article complements the ongoing and (so far) mostly theoretically oriented research into the impact of the GDPR on organizations, examining possible factors that play a key role in the implementation of the Regulation. This article first provides empirically-based identification of the factors that determine the cost of the process itself and responds to the existing research gap. For managerial practice, this article is the basis for budgeting implementation projects and points out areas that need to be given special attention. This research has a higher value for companies with up to 10 employees, which make up majority of entities on the market.

Several limitations of this research can be identified, while the limitations are closely related to the authors' proposals for further research. At first, this article dealt with the study of organizations operating in the Czech Republic. Further research in this regard should also address the examination of key factors in the context of other, especially the EU, countries. The second limitation of this research are the factors that have been implied from previously published works. As partially outlined in the results of this study, other factors and possible moderating variables need to be considered in future research.

Finally, further research should focus on other characteristics of the implementation of GDPR such as technological solutions used or quality of the implementation.

Funding

This work was supported by the internal grant No. SGS-2021-017 of the University of West Bohemia.

Author contributions

This article was compiled in collaboration of two authors – Adam Faifr and Martin Januška. Adam Faifr was responsible for the research design and literature review. Martin Januška was responsible for conducting the survey within the organizations and was responsible for subsequent data analysis. All the remaining aspects of this work were processed in close cooperation of both authors.

Disclosure statement

We declare to have no conflict of interest with other parties.

Acknowledgements

The authors would like to take this opportunity to thank Mrs. Markéta Maňourová for her participation in creating and processing the online survey, as well as for the subsequent data collection. The authors are also grateful to anonymous reviewers for their comments, which contributed to the final version of the article.

References

Almeida Teixeira, G., Mira da Silva, M., & Pereira, R. (2019). The critical success factors of GDPR implementation – a systematic literature review. *Digital Policy, Regulation and Governance*, *21*(4), 402–418. https://doi.org/10.1108/DPRG-01-2019-0007

Beckett, P. (2017). GDPR compliance: Your tech department's next big opportunity. *Computer Fraud & Security*, *2017*(5), 9–13. https://doi.org/10.1016/S1361-3723(17)30041-6

Bleier, A., Goldfarb, A., & Tucker, C. (2020). Consumer privacy and the future of data-based Innovation and marketing. *International Journal of Research in Marketing*, *37*(3), 466–480. https://doi.org/10.1016/j.ijresmar.2020.03.006

Council of the European Union. (2013). *Commission Recommendation of 6 May 2003 concerning the definition of micro, small and medium-sized enterprises*. https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2003:124:0036:0041:EN:PDF

Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). SAGE Publications, Inc. https://upog.pw/lixez_hibuk_ky_ke_letir.pdf

Czech Chamber of Commerce. (2018). *Účet za GDPR? Podnikatele nařízení vyjde na 25 miliard korun*. Retrieved April 8, 2020, from https://www.komora.cz/press_release/ucet-za-gdpr-podnikatele-narizeni-vyjde-na-25-miliard-korun

Datoo, A. (2018). Data in the post-GDPR world. *Computer Fraud & Security, 2018*(9), 17–18. https://doi.org/10.1016/S1361-3723(18)30088-5

Diamantopoulou, V., Tsohou, A., & Karyda, M. (2019). General Data Protection Regulation and ISO/IEC 27001:2013: Synergies of activities towards organisations' compliance. In *Lecture notes in computer science: Vol. 11711. Trust, privacy and security in digital business* (pp. 94–109). Springer Publishing. https://doi.org/10.1007/978-3-030-27813-7_7

Diamantopoulou, V., Tsohou, A., & Karyda, M. (2020). From ISO/IEC27001:2013 and ISO/IEC27002:2013 to GDPR compliance controls. *Information and Computer Security*, *28*(4), 645–662. https://doi.org/10.1108/ICS-01-2020-0004

European Parliament, & Council of the European Union. (2016). *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)*. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679&from=EN

Everett, C. (2011). Is ISO 27001 worth it? *Computer Fraud & Security*, *2011*(1), 5–7. https://doi.org/10.1016/S1361-3723(11)70005-7

Garber, J. (2018). GDPR – compliance nightmare or business opportunity. *Computer Fraud & Security*, *2018*(6), 14–15. https://doi.org/10.1016/S1361-3723(18)30055-1

Gal, M. S., & Aviv, O. (2020). The competitive effects of the GDPR. *Journal of Competition Law & Economics*, *16*(3), 349–391. https://doi.org/10.1093/joclec/nhaa012

Hofman, D., Lemieux V. L., & Batista, D. (2019). The margin between the edge of the world and infinite possibility: Blockchain, GDPR and information governance. *Records Management Journal*, *29*(1/2), 240–257. https://doi.org/10.1108/RMJ-12-2018-0045

Hoofnagle, C. J., Sloot, B., & Borgesius, F. Z. (2019). The European Union general data protection regulation: What it is and what it means. *Information & Communications Technology Law*, *28*(1), 65–98. https://doi.org/10.1080/13600834.2019.1573501

Huber-Carol, C., Balakrishnan, N., Nikulin, M. S., & Mesbah, M. (2002). *Goodness-of-fit tests and model validity.* Springer Publishing. https://doi.org/10.1007/978-1-4612-0103-8

Khan, J. (2018). The need for continuous compliance. *Network Security*, *2018*(6), 14–15. https://doi.org/10.1016/S1353-4858(18)30057-6

Kindt, E. J. (2018). Having yes, using no? About the new legal regime for biometric data. *Computer Law & Security Review*, *34*(3), 523–538. https://doi.org/10.1016/j.clsr.2017.11.004

Kounoudes, A. D., & Kapitsaki, G. M. (2020). A mapping of IoT user-centric privacy preserving approaches to the GDPR. *Internet of Things*, *11*, 100179. https://doi.org/10.1016/j.iot.2020.100179

Larrucea, X., Moffie, M., Asaf, S., & Santamaria, I. (2020). Towards a GDPR compliant way to secure European cross border Healthcare Industry 4.0. *Computer Standards & Interfaces*, *69*, 103408. https://doi.org/10.1016/j.csi.2019.103408

Lindgren, P. (2018). GDPR regulation impact on different business models and businesses. *Journal of Multi Business Model Innovation and Technology*, *4*(3), 241–254. https://doi.org/10.13052/jmbmit2245-456X.434

Longras, A., Pereira, T., Carneiro, P., & Pinto, P. (2018). On the track of ISO/IEC 27001:2013 implementation difficulties in Portuguese organizations. In *2018 International Conference on Intelligent Systems* (pp. 886–890). IEEE. https://doi.org/10.1109/IS.2018.8710558

Maňourová, M. (2019). *GDPR – Evaluation of the impacts of GDPR on businesses in the Czech Republic*. University of West Bohemia, Pilsen, Czech Republic. https://dspace5.zcu.cz/handle/11025/38705

Martin, K. D., Kim, J. J., Palmatier, R. W., Steinhoff, L., Stewart, D. W., Walker, B. A., Wang, Y., & Weaven, S. K. (2020). Data privacy in retail. *Journal of Retailing*, *96*(4), 474–489. https://doi.org/10.1016/j.jretai.2020.08.003

McCall, B. (2018). What does the GDPR mean for the medical community? *The Lancet, 391*(10127), 1249–1250. https://doi.org/10.1016/S0140-6736(18)30739-6

Mesquida, A. L., & Mas, A. (2015). Implementing information security best practices on software lifecycle processes: The *ISO/IEC 15504 Security Extension*. *Computers & Security*, *48*, 19–34. https://doi.org/10.1016/j.cose.2014.09.003

Nguyen, L. D., Le-Hoai, L., Tran, D. Q., Dang, C. N., & Nguyen, C. V. (2019). Effect of project complexity on cost and schedule performance in transportation projects. *Construction Management and Economics*, *37*(7), 384–399. https://doi.org/10.1080/01446193.2018.1532592

Nonnemann, F. (2011). *Personal data protection during information providing by public organizations*. Ministry of the Inferior of the Czech Republic. Retrieved April 8, 2020, from https://www.mvcr.cz/clanek/clanek/ochrana-osobnich-udaju-pri-poskytovani-informaci-verejnou-instituci.aspx

Park, M., Choi, S., Shin A. M., & Koo, C. (2013). Analysis of the characteristics of the older adults with depression using data mining decision tree analysis. *Journal of Korean Academy of Nursing*, *43*(1), 1–10. https://doi.org/10.4040/jkan.2013.43.1.1

Parliament of the Czech Republic. (2019). *ZÁKON ze dne 12. března 2019 o zpracování osobních údajů*. https://eur-lex.europa.eu/legal-content/CS/TXT/PDF/?uri=NIM:272327

Perry, R. (2019). GDPR – project or permanent reality? *Computer Fraud & Security, 2019*(1), 9–11. https://doi.org/10.1016/S1361-3723(19)30007-7

Prakash, M., & Singaravel, G. (2015). An approach for prevention of privacy breach and information leakage in sensitive data mining. *Computers & Electrical Engineering, 45*, 134–140. https://doi.org/10.1016/j.compeleceng.2015.01.016

Quinn, O., & Quinn, L. (2018). Big genetic data and its big data protection challenges. *Computer Law & Security Review, 34*(5), 1000–1018. https://doi.org/10.1016/j.clsr.2018.05.028

Sirkin, M. R. (2006). The Chi-Square test, statistics for the social sciences. In Sirkin, M. R., *Statistics for the Social Sciences* (3rd ed.). SAGE Publications, Inc. https://doi.org/10.4135/9781412985987.n12

Sirur, S., Nurse, J., & Webb, H. (2018). Are we there yet? Understanding the challenges faced in complying with the General Data Protection Regulation (GDPR). In *25th ACM Conference on Computer and Communication Security* (pp. 88–95). Canada. https://dl.acm.org/doi/10.1145/3267357.3267368

Starčevič, K., Crnkovič, B., & Glavaš, J. (2018). Implementation of the General Data Protection Regulation in companies in the Republic of Croatia. *Ekonomski Vjesnik / Econviews*, *31*(1), 163–176. https://pdfs.semanticscholar.org/d75a/1a38e0a560f7ac9dde52c33a387c0c6fe21a.pdf

Strickland, J. (2016). *Data analytics using open-source tools* (1st ed.). Lulu.com.

Sue, V. M., & Ritter, L. A. (2007). *Conducting online surveys*. SAGE Publications, Inc. https://doi.org/10.4135/9781412983754

Tamburri, D. A. (2020). Design principles for the General Data Protection Regulation (GDPR): A formal concept analysis and its evaluation. *Information Systems*, *91*, 101469. https://doi.org/10.1016/j.is.2019.101469

Tankard, C. (2016). What the GDPR means for businesses. *Network Security*, *2016*(6), 5–8. https://doi.org/10.1016/s1353-4858(16)30056-3

The office for personal data protection. (2018). *S účinností GDPR končí oznamovací povinnost správců.* https://www.uoou.cz/s-ucinnosti-gdpr-konci-oznamovaci-povinnost-spravcu/d-28855

Tikkinen-Piri, C., Rohunen, A., & Markkula, J. (2018). EU General Data Protection Regulation: Changes and implications for personal data collecting companies. *Computer Law & Security Review*, *34*(1), 134–153. https://doi.org/10.1016/j.clsr.2017.05.015

Udroiu, M., Dumitrache, M., Sandu, I., & Brezulianu, A. (2018). Implementing an integrated information system designed for Romanian public entities. *Studies in Informatics and Control*, *27*(3), 369–376. https://doi.org/10.24846/v27i3y201812

Yuan, B., & Li, J. (2019). The policy effect of the General Data Protection Regulation (GDPR) on the digital public health sector in the European Union: An empirical investigation. *International Journal of Environmental Research and Public Health*, *16*(6), 1070. https://doi.org/10.3390/ijerph16061070