Identifying Emerging Challenges for ICT industry in Ireland: Multiple Case Study Analysis of Data Privacy Breaches

Emergent Research Forum (ERF)

Plamen Petkov
ADAPT Centre, Dublin City University
plamen.petkov@adaptcentre.ie

Markus Helfert
Dublin City University
markus.helfert@dcu.ie

Abstract

With a highly creative and talented workforce, an open economy and a competitive corporate tax environment, Ireland had successfully attracted top global information technology firms. However, with the introduction of the new policies and regulation posed by the General Data Protection Regulation (GDPR), companies are facing new challenges in terms of tougher penalties, stricter internal policies, and privacy training programs. In this emerging research, we present a pilot study and initial analysis of scenarios in Irish IT industry sector where infringement of the data protection laws occurred. The aim of the analysis is to identify directions and trends with respect to the type of breach/disclosure, where the breach occurred, who was involved in the process, etc. We anticipate that the outcomes of this study assist defining clearer requirements for IT companies in Ireland in relation to the GDPR.

Keywords

General Data Protection Regulation, Irish IT sector, Multiple Case Study, Industry Requirements

Introduction

Ireland is a small highly globalized economy, with a large exporting sector, and a significant number of multinational corporations. Ireland has one of the lowest corporate tax rates in Europe – 12.5% compared with countries such as the UK where that number is near twice as higher – 20% or counters such as Germany Belgium, France, and United states of America where the corporate tax rates pass the 30th mark of the percentile (PwC, 2016). On a separate note, Ireland has one of the most educated workforces in the world - according to the Organization for Economic Co-operation and Development (OECD) - 52% of 25-34 year-olds have a third level qualification which 10% higher than the OECD average (OECD, 2015). Considering statements above, plus the fact that after the Brexit, Ireland will effectively remain the only native English speaking country member of the European Union, makes the country extremely attractive for international business, especially within the area of Information and communications technology (ICT) services. In the past decade, Ireland has become a global technology hub attracting the strategic business activities of big ICT companies such as Amazon, LinkedIn, PayPal, eBay, HP, Facebook, Apple etc. Ireland is the also the European data center location of choice for world leaders including IBM, Microsoft, Google, Yahoo, MSN and Adobe and is now anticipated to become a global cloud center of excellence.

On the other side, the latest technological advancements worldwide had raised numerous ethical issues related to the individual rights and the usage of personal information. More particularly, increased number of security breaches and leakages of personal information is a growing concern for individuals’ privacy and companies which use personal data. To bridge the gap between the way technologies are utilized and the human rights, ethicist, EU policy makers, trade unions and IT specialists had developed large set of complex privacy rules and regulations, stacked under the umbrella of the General Data Protection Regulation (GDPR) legislative paper which is scheduled to enter into force in May 2018. The GDPR sets new rules to the data controllers with respects to the handling of person information and more specifically about the way user’s consent is obtained. It also introduces new obligations such as privacy by
design, the necessity of Data Protection Officer (DPO) whose purpose is to ensure proper compliance and prompt reporting of security and data breaches to the Office of the Data Protection Commissioner (ODPC, 2017). Last but not least GDPR brings changes in terms of penalties associated with non-compliance – e.g. breaching the policy rules can result in fines that span up to €1 million euro or 4% of a company’s recurring revenue (EU, 2016).

Logically, every IT organization should strive to prevent the improper use or disclosure of personal, customer, or employee information. Irish Computer Society claimed that three out of four companies had already appointed DPOs, however with the level of employer’s and staff confidence with regards to data security and privacy continues to plummet down as the new GDPR emerges on the horizon (ISC, 2016). This could be due to lack of proper training programs among employees, or failure to grasp the entirety of the data flows within organizations or simply due to inability to detect poorly designed business processes which could potentially contribute to the disclosure personal information.

In this paper, we present a pilot study and initial analysis of scenarios in Irish IT industry sector where infringement of the data protection laws occurred. The analysis is based on data privacy breach cases reported to the ODPC in Ireland. The aim of the analysis is to identify leads and trends with respect to the type of breach/disclosure, where the breach occurred; who was involved in the process of disclosing information, what is was the official statement by the appointed commissioner's staff in relation to the case, etc. Ultimately, we anticipate that the outcome of this analysis leads to building knowledge which could be potentially used for educational and training purposes as well as a lead of defining clearer requirements for IT companies in Ireland in relation to the GDPR.

**Research Methodology**

Given the exploratory nature of this work, conducting multiple case study was found to be the most suitable approach at this stage of the project. Yin (2013) defines the scope of a case study research as an “empirical inquiry that investigates a contemporary phenomenon within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident”. In this project, we perform analysis of a number of cases to gain understanding and explore boundaries of systems, (i.e., composition of processes) in which disclosure/breach of privacy occurred including the units such as technologies, staff, third party organizations, etc. that lead to the breach. Next paragraphs briefly describe the rationale behind case selection, data collection, and analysis.

**Selecting cases:** The proper selection of cases is considered crucial by all the case study methodologists (Stake, 2006) (Yin, 2013) (Walsham, 1995). As a consequence, criteria need to be clearly established to select the most suitable cases. In this initial study, cases were selected based on the following requirements: 1) case has to describe, in an observable way, a breach of individual’s privacy within the IT sector, 2) case has to be accountable and accessible, and 3) case has to date no longer than 2013. After defining the requirements that have to be employed during the selection of the cases, another decision that had to be taken was in relation to the number and sources of cases to be studied. Yin (2013) recommended collecting data from multiple sources to reduce bias. However, in this pilot work, we consider ODPC to be the ultimate source of truth as it is the ultimate body to exercise legislative regulations. As to the number, 20 cases were selected to be analyzed as we consider this would provide sufficient output at this stage of the research.

**Data collection:** Case studies were represented in a form of free text structured in several paragraphs. Each case had a size of approximately 500 – 1000 words, summarized by an ODPC’s official. With respect to the structure, each case contained an introductory paragraph which provides an overview of the case and its context followed by the description of the issue(s) raised by the complainant to the commissioner’s office. Then, the next section describes procedures and actions taken by DPC to conduct an investigation on behalf of the complainant. The procedures involve relaying complainant’s, concern and seeking a formal response from the data controller (e.g. employer or a person accountable for the fair and secure processing and storing of complainant’s, information). During the investigation process, DPC may also require from the data controller to provide a proof of compliance with the data protection act, in terms of evidence such as support documents describing the purpose(s) of the personal data being (or not) misused/breached. Having the request clearly posed by the DPC, the third section of the text sought to present the position of the of the data controller in regards with presented complaint. Considering the
pieces of evidence provided by the data controllers in relation to the raised concern, last part of each case study describes the decision taken by the DPC with respect to the violated legislative statement (if any) of the data protection acts as well as how this decision affect the outcome of the case. In many cases, the last section of the summary also includes recommendations and best practices for improvement to avoid future organizational disappointments and infringements of the data protection law.

**Analysis:** According Neumann (2005) the process of data analysis involves “a search for patterns in data”. Once a pattern identified, the qualitative researcher moves from the description event to a more general interpretation of its meaning. This is consistent with the ultimate goal of inductive case studies, which is “to uncover patterns, determine meanings, construct conclusions” (Neuman, 2005, p.426). Additionally, Patton (1990) argues that the objective of analyzing the data from documents (i.e. the main data sources for this study) is to identify codes, categories, and patterns that are encapsulated in the data. All in all, coding is found suitable in relation to the interpretive nature of this study. In fact, “coding facilitates the organization, retrieval, and interpretation of data and leads to conclusions on the basis of that interpretation” (Lockyer, 2004). Considering the introductory stage and limited research resources of this research, the coding was conducted by the main author of this manuscript. However, in future, we will involve multiple (at least 2 additional) researchers to increase the objectiveness of the codifying process by analyzing cases independently of each another and comparing their results. The method of triangulation will be used in cases of disagreements between researchers. The coding criteria as shown in Table 1 was based on the initial objective of this study – to identify trends in the type of the privacy breaches related to the processes and/or people caused the issue as well as link the new GDPR principle and stages of data governance.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case name</td>
<td>Name of the case; Normally indicates the context of the observed case</td>
</tr>
<tr>
<td>Complaint</td>
<td>Short description of complaint, e.g. unsolicited marketing SMS campaign</td>
</tr>
<tr>
<td>Type breach/disclosure</td>
<td>Indicates the type of breach or disclosure, e.g. unfair processing of data</td>
</tr>
<tr>
<td>Process</td>
<td>Briefly describes the process that leads to the violation, e.g. broken opt-out link to a service</td>
</tr>
<tr>
<td>GDPR principle</td>
<td>Violated GDPR principle(s), e.g. data purpose limitation</td>
</tr>
<tr>
<td>Data lifecycle reference</td>
<td>The stage of the data lifecycle during which a disclosure occurred, e.g. obtaining customer data unlawfully</td>
</tr>
<tr>
<td>Outcome</td>
<td>Briefly describes the decision taken by the DPC, e.g. prosecution</td>
</tr>
<tr>
<td>Recommendation</td>
<td>Indicates recommendation/best practices to the data controller</td>
</tr>
</tbody>
</table>

**Table 1. Criteria for Case Study Analysis**

**Discussion and Findings**

The ODPC provided us with a database of above 200 cases distributed across different industry domains from the past 15 years. For the purposes of this pilot study, we have chosen an initial number of 20 cases to be analyzed that were positioned partly or entirely within IT domain. Many of those cases, however, covered other industries such as utility providers, government services, healthcare, as well as retail and marketing. With respect to the allegations filed by complainants, our analysis indicated a great diversity of cases - from unauthorized access to customer information by employees, through illegal sharing of medical-related information with insurance companies, to the unfair processing of personal information by government authorities, and unsolicited marketing campaigns. Figure 1 displays some important remarks in relation to the GDPR principles, and stages of the data lifecycle a breach of the data protection law occurred. Our initial analysis showed the primary reason for infringement of individual’s rights was the unfair and unlawful processing of personal information - 9 out of 20 cases. This is somehow logical since in order to conduct fair and transparent processing data controllers have to follow other principles posed by the GDPR such as purpose limitation (e.g. clearly describe the purpose and lawfully obtain consent for data being processed), minimization (data has to be limited to what is necessary for relation
Multiple Case Study Analysis of Data Privacy Breaches

to the purposes), strictly defined retention periods (personal data should be stored for no longer than is necessary for the purposes for which it is processed), security and accuracy. Accountability is a key principle, necessary to prove the transparent processing of personal information. In 8 of the cases, the data controller failed to provide solid proof of compliance which is also one of the reasons why a big number of cases resulted in prosecution. In fact, data controllers in 14 out of 20 cases ended up being prosecuted partly because of failure to provided accountability. This indicates the strictness and seriousness of the DPC in relation to abilities of data controllers to demonstrate, compliance with the data protection principles. Interestingly enough, only in 3 of the cases, a security breach within the information system was accountable for personal data disclosure.

The second diagram as shown in Figure 1, b) depicts the distribution of cases in which individuals’ privacy rights were breached in relation to the stage of the data lifecycle during which the breach occurred. The difference between primary and secondary data shown on the charts is that the former is obtained through direct observations, e.g. data lifecycle stage during which breach occurred was explicitly reported by the ODPC official, while latter is obtained through indirect analysis, e.g. data maintenance involves processes of manipulating data such as retrieving, updating and storing. Our direct observation indicated that breaches happened nearly throughout the entire lifecycle with an increased number of breaches during the storing (e.g. improper storing procedures) stage and applying stage (e.g. personal information was not used for the purposes the user agreed to in the terms and condition statement). Another practical example is a failure to maintain an opt-out link which led to unsolicited marketing SMS campaign, or in yet another, a mistake by a staff member to properly update customers’ records in all databases within organization resulted in excessive email spam. A few other cases resulted in non-compliance with the retention periods by failing to maintain databases containing records of former customers. Last but not least, we were able to identify the following reason why breaches occurred: human error – 9/20, poorly designed internal processes or system failure – 4/20, unawareness of data flows/bad data quality – 3/20; and 4 of the cases were labeled with a reason different from ones mentioned above. Despite our initial observation, we recommend further empirical analysis, to determine correlations between different breach causalities.

**Figure 1. Number of Cases in Relation to: a) Violated GDPR Principle; b) Stage of Data Lifecycle During which Violation Occurred;**

Conclusion and Future Work

Ireland is the second largest exporter of computer and IT services in the world. With a highly creative and talented workforce, an open economy, and a competitive corporate tax environment, Ireland has
Multiple Case Study Analysis of Data Privacy Breaches

successfully attracted top global information technology firms. However, with the introduction of the new GDPR, companies are facing new challenges in terms of tougher fines, stricter internal policies, solid privacy training programs, a need for increased awareness of internal and external business processes and better data governance. In fact, according to this pilot study, nearly 30% of the cases resulted in data privacy breaches due to unawareness of the data flows or poorly composed/optimized business and IT processes. Particularly, this is an alarming trend since the effectiveness of any privacy training programs would highly depend on the accuracy and transparency of system/organizational processes. Hence, Irish practitioners would need comprehensive tools for documenting, analyzing and optimizing their business processes to ensure data quality, accountability, transparency, and compliance with the GDPR. This would be specifically useful for cases when flows of personal data scope countries outside Europe.

At this stage, we have analyzed a poll of 20 case studies provided by the ODPC of Ireland. Our future work will involve analyzing more scenarios of data breaches to derive more accurate results and extend the knowledge base. This process will involve a closer cooperation with the ODPC, in terms of providing access to data related to cases where data privacy laws were infringed. Learning from the mistakes of others is one of the most effective ways to clearly define requirements, ensure conformance and stay on the top of the competition. In a longer term, we plan to extend this framework to countries outside Ireland and build a worldwide accessible database which will contribute to industries operating with personal information of European citizens. The database will contain an analysis of cases, with a particular focus on the section of the workflow where the breach occurred. This also will include representing information about the particular personal data type that was breached as well as factors and actors contributed to this breach. To represent this transparently we aim to employ the Business Process Modeling Notation specification which is widely recognized standard for process modeling (White, 2004). The necessity for all stakeholders to fully understand the mapping of the process to the data flow is critical for the successful compliance with the GDPR. Going beyond the simplistic description of breach cases provided by the ODPC’s official web page, delivering a detailed and transplant analysis of data flows will substantially increase compliance and understanding of GDPR among communities which comprise not only data protection officers but also system architects, developers, data analysts as well as strategic and management members.

Acknowledgements

This study was supported by the ADAPT research center at Dublin City University, and the Data Protection Commission of Ireland.

REFERENCES