Approaching the Data Protection Impact Assessment as a legal methodology for the evaluation of the level of privacy by design achieved by technological proposals. A special reference to Identity Management systems

ABSTRACT: The process of digitalization of societies and innovation is involving the fast introduction of new technologies in different sectors. However, the development of technology represents a challenge as it involves technical, legal, economic and social aspects that have to be considered since its conception or design. The aim of this paper is to offer an adaptation of an existing legal methodology, the Data Protection Impact Assessment, as a legal obligation to evaluate technological proposals and assure compliance with privacy by design requirements. For that purpose, we will refer to the specific case of Identity Management technologies. We introduce We introduce the main challenges in the sector of Digital Identity Management as well as the importance of covering the "architecture" and "user" sides in the development of safer technologies by introducing concrete examples. Finally, in order to provide a more practical vision of the methodology to adapt the Data Protection Impact Assessment, we refer to the work developed in the research project OLYMPUS in the evaluation of its privacy implications. By introducing this example, the paper offers a specific methodology directly reusable for the study of technological proposals in IdM but that can be adapted to any other sector.

Keywords: Data Protection Impact Assessment, Privacy by design, Identity Management, GDPR

1 INTRODUCTION

The Identification of individuals in online environments plays a key role in the guarantee of safe and trustworthy online activities. Digital identification and subsequent authentication have become the vehicle for the development of online environments and services where it was necessary to ascertain that the right individual is behind the "screen" and therefore has the corresponding right to perform that specific action or process. However, digital identification and authentication of individuals is a complex task and is posing a set of risk that threatens the basis of free and democratic societies [1].

On the other hand, the change in the habits and communication means have made of identity theft one of the most widespread forms of cybercrime [2,3,4]. Identity theft is a specific type of cybercrime linked to digital identity and it can be defined in different ways. Following the definition provided by the U.S. Department of Justice "identity theft and identity fraud are terms used to refer to all types of crime in which someone wrongfully obtains and uses another person's personal data in some way that involves fraud or deception, typically for economic gain" [5]. In some cases, identity theft is included as a category of identity fraud, or identity theft is considered as a precursor of identity fraud [6]. Consequently, identity theft does not appear as a stand-alone crime and its consequences can affect the victims for a long time [7]. Furthermore, as stated with regard to surveillance practices, in the case of identity theft, the excessive "centralization" in Identity Management (hereafter, IdM) in a single, or a reduced number of Identity Providers (hereafter, IdP) in the case of federated IdM, favours targeted attacks to thereof as they represent a single point of failure.

The fight against surveillance practices and identity theft requires a multidisciplinary approach involving the collaboration of technical, legal and social experts in order to assess the problem, propose solutions and in a last stage, validate them. To limit surveillance practices, new forms of safe and privacy-preserving IdM will have to be designed and Regulations will have to support the adoption of such technologies.

With regard to the prevention of identity theft, the adoption of safe technologies is essential. Nevertheless, safe technologies do not only refer to complex techniques or encryption processes, but they also imply comprehensible concepts for end users, who will also be decisive in the fight against this specific form of cybercrime. The evaluation of the level of privacy achieved in a technology's design, and its balance with security and usability requirements can be studied by means of different methodologies. In this sense, the methodology par excellence for the study of privacy implications is the Data Protection Impact Assessment (hereafter, DPIA), envisaged by the Article 35 of the General Data Protection Regulation (hereafter, GDPR), among others. In addition, other methodologies such as the risk analysis (usually subsumed in the DPIA) or new approaches, such as the concept of privacy engineering adopted by the Spanish Data Protection Agency are of extreme relevance.

The aim of this paper is to offer a legal methodology for the study of the level of privacy by design achieved by technological proposals prior their implementation. For that purpose, we will make reference to the two "sides" or "dimensions" that must be covered in the development of privacy-preserving technologies, including specific examples in the sector of IdM. Therefore, a comprehensive explanation of our proposal of adaptation of the DPIA to the scenario above-mentioned (technologies that have not been implemented yet) will be provided, specifying phases, steps and special considerations that must be taken into account in these cases to facilitate subsequent analysis in context-specific scenarios.

2 PRIVACY-PRESERVING TECHNOLOGIES IN IDM

IdM technologies are evolving through the development of new architectures and the implementation of encryption techniques that face the challenges referred above and that ensure a better compliance with privacy by design requirements. Privacy by design has been envisaged as a legal obligation in Article 25.1 of the GDPR. To determine the content of this obligation we should refer to Article 5 of the same legal text that establishes the principles that all data processing activity must fulfil (i.e., lawfulness, data minimization and security in the data processing) and that must also be taken into account in the design of a technology. Along this process, technical experts have remarked the need to design more than resilient architecture but also to protect the user thereof [8]. In this sense, we have distinguished a double direction in which IdM technologies must be improved and where innovations must be complemented and supported between them. These dimensions refer, on the one hand, to the architecture or the "internal side", and, on the other hand, to the user or the "external side".

The design of strong architectures has been one of the main objectives in the improvement of IdM. In this sense, important research has been carried out in the area. By way of example, we make reference to the European Union research project OLYMPUS [9]. OLYMPUS is an IdM system developed in the framework of delegated/ federated IdM. More specifically, it introduces three main innovations [10]:

a) It distributes the task of the IdP among several IdPs (integrating the virtual IdP) by means of novel cryptographic approaches that allows the password “fragmentation”.

b) It envisages a possibility of offline deployment through Privacy Attribute-Based Credentials (hereafter, p-ABCs) cryptographic techniques.

c) It allows the redistribution of fragments of passwords among the partial IdPs in established periods of time.

The result of these innovations is translated into two main possibilities. On the one hand, the possibility to prevent surveillance practices by means of the implementation of p-ABCs cryptography in offline scenarios. In other words, the user will be able to request the issuance of a credential containing a set of attributes (e.g., the name, data of birth…), that he/she will store in his/her digital wallet and will use to authenticate in a later stage. By making use of this process, the connection between the IdP and the service provider is “broken”.

On the other hand, OLYMPUS’ innovations improve the prevention of identity theft. Indeed, the distributed architecture introduced by OLYMPUS hampers different types of identity theft. Regarding token identity theft, for the token issuance OLYMPUS requires the collaboration of all the partial IdPs which conform the vIdP, demanding the attacker to have the control over all the structure, as user’s password (necessary for the token issuance) appears disaggregated through thereof. In addition, against traditional identity theft attacks or those that relate to the discovery of a fragment of password, OLYMPUS allows the redistribution of the segments of password through the mechanism of “key-resharing”, posing countless scenarios to the potential attacker.

Nevertheless, measures to prevent identity theft attacks do not only require the development of resilient architectures but must also focus their attention on the user. As stated in the introduction, identity theft has become one of the most widespread forms of cybercrime and the means for its commission have evolved, representing social engineering techniques one of the most common means for its commission [11]. From the perspective of technical proposals in this section we could refer to The Expanded Password System (hereafter, EPS).

The EPS consists in an authentication method that introduces the possibility of converting text passwords into images. The authentication process takes place by selecting a set of images that only the user is able to select correctly since these images are associated with his/her autobiographical or episode memories. Nevertheless, this collection of images will be exclusively presented to the user, since software will translate the images into text passwords, that will be the ones finally stored [12,13]. In such scenario it will be more difficult to steal user’s password and he/she will easily detect attempts of fake logging since the images selected by him/herself would not appear. Indeed, as Hitoshi Kokumai notes, “a would-be phisher can easily copy the log-in screen and show it to a target user whose User ID is known. But the phisher does not know which image was registered by the user as the credential of the genuine log-in server as against the other images, whereas both the user and the genuine log-in server know which one was registered” [14]. Consequently, if the user is given a password box or the choices do not include the registered images, the user would know immediately that it is an attempt of phishing.

Besides, there exist other mechanisms such as Firefox Monitor or Google leaked-password checker to detect possible privacy violations in an early stage and minimize the possible harm of online risks [15]. Data Protection Authorities and public/private intermediaries develop a labor of concerning individuals in choosing strong passwords [16] as well as with regard to the understanding of social engineering techniques and how to avoid becoming victims of cybercrime. Nevertheless, these have just been a few examples of how to implement privacy-preserving mechanisms that improve IdM and face the challenges that it is raising nowadays. All these technical “tools” have required a previous evaluation that determines their suitability and their compliance with regulations. From our side, as legal experts, the challenge is how to determine whether a specific technology fulfils privacy and data protection requirements prior its implementation. For that purpose, what we propose is the adaptation of the DPIA methodology. The DPIA is not a new tool, but the innovative approach given to it in this paper differs from traditional uses. Indeed, we propose in the following section the possibility of using the methodology of the DPIA to perform a preliminary assessment over technological proposals prior their implementation, in order to ease further context-based studies, and above all, assure compliance with privacy of design requirements.

3 THE DPIA AS A METHODOLOGY TO ENSURE GDPR COMPLIANCE IN TECHNOLOGICAL PROPOSALS

The importance of considering privacy as part of a system’s development process is widely accepted as an essential aspect towards the development of privacy-aware systems [17]. In addition, there exist an effort in the development and standardization of technological solutions that are privacy friendly. In such scenario, Privacy Enhancing Technologies (PETs), defined in the Communication of European Commission to the European Parliament and Council as “a coherent system of ICT measures that protects privacy by eliminating or reducing personal data or by preventing unnecessary and/or undesired processing of personal data, all without losing the functionality of the information system” [18] reach an extraordinary importance.

Nevertheless, the development of privacy respectful technologies is a complex task. Although technical experts manage some privacy concepts, they tend to omit essential legal requirements that could be decisive in the final evaluation of a technology or that imply, in some cases, an inadequate approach to the technology design. The Spanish Data Protection Agency has referred to the concept of privacy engineering [19] as the process for the implementation of privacy in the lifecycle of those information systems where the processing of personal data takes place. In this process, we distinguish a set of phases. The first step will take place even before the design of the technology had started. In this sense, it is critical that the properties and functionalities that a system must fulfill in terms of privacy are clearly and previously stated. These properties must refer to the minimal requirements that would make possible the implementation of the technology. The second step will consist in the design of the architecture and implementation of the elements in the system that cover the privacy requirements previously defined. Finally, it must be confirmed that privacy requirements have been correctly implemented and satisfy expectations and needs. In the final evaluation, or during the technology’s design, modifications or safeguards can be proposed to integrate the final architecture design.

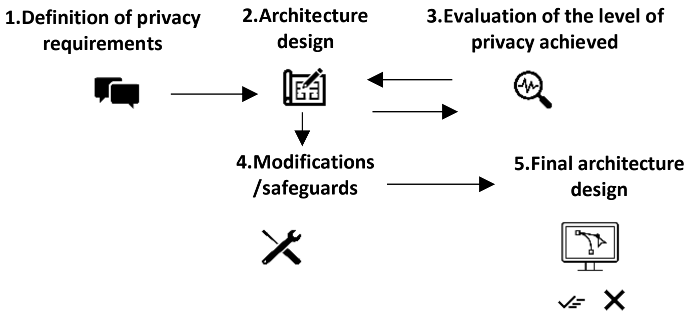
1.Definition of privacy requirements
2.Architecture design
3.Evaluation of the level of privacy achieved
4.Modifications /safeguards
5.Final architecture design

Figure 1: Privacy engineering phases [Source: Authors, 2021]

Consequently, these prior evaluations or, in other words, collaborative designs can bring important benefits and enhance the possibilities of success of a technology. Nevertheless, developing a methodology for the evaluation of technological proposals prior their implementation represents a challenge as it involves not only technical but very relevant legal concepts as well. The DPIA is an instrument specifically envisaged for the assessment of data processing operations in those cases where, as stated in Recital 84 of the GDPR,“processing operations are likely to result in a high risk to the rights and freedoms of natural person” [20], and as Article 35 adds “where a type of processing in particular using new technologies, and taking into account the nature, scope, context and purposes of the processing, is likely to result in a high risk” [20]. From these two articles we can extract that the DPIA is a tool specifically envisaged for context-based scenarios. However, some of the privacy and data protection risks detected in a DPIA are linked to the technology deployed. In this sense, to perform this study in a prior stage could prevent from discarding technological proposals with big potential, or in the worst cases, implementing technological solutions that imply a high risk but that are too expensive to modify at that phase of development. For these reasons, we propose the adaptation of the DPIA for the appraisal of technological solutions prior their implementation by giving in a new approach to the DPIA methodology in “layers”. By way of example, we will propose a methodology by taking as reference the work developed in the research project OLYMPUS.

3.1 First layer

The first step in the development of a DPIA is the description of the scope of study. This description should content at least an explanation about the technological proposal and its main differences or improvements concerning existing technologies. Explained the characteristics of the technology, the data flows must be described. The data flows refer to the transfer, exchange, storage or modification of data that takes place in the use of the technology or the performance of operations involving the processing of personal data. In the case of IdM systems the data flows refer to the lifecycle of digital identities, that is to say, how enrollment, authentication and digital identity management will be carried out in this specific system. Furthermore, it would also be recommendable to include a graphical representation of the architecture design to make the technology easily understandable and to highlight the main differences introduced. This phase does not represent major difficulties, but the description of the technology must be made using a language that can be understood by technical and non-technical experts so that compliance with legal requirements is properly considered.

Once described the technology and the data flows, before proceeding in the DPIA, it would be recommendable to analyze at least one prior issue. Considering the latest developments in cryptography and the emergence of techniques aiming to reach anonymity of the data, the first aspect to study is whether the data processed can actually be considered as personal data and, therefore, if the GDPR is applicable. If from this step, we conclude that the data processed do not qualify as personal data the assessment would finish in this phase.

In order to determine whether the data processed can be qualified as personal data, we have to refer to the definition contained in Article 4(1) of the GDPR which states that “personal data means any information relating to an identified or identifiable natural person (data subject)” [20]. Hence, personal data is the information that directly or indirectly relates to an identified or identifiable natural person. Conversely, when data does not relate to an identified or identifiable natural person, data must be considered anonymous, which according to what is stated in Recital 26 of the GDPR does not fall under the scope of the principles of data protection and our DPIA will conclude at this phase.

Consequently, personal data does not only refer to the data that directly identify an individual, but that make individuals identifiable [21]. An identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person. In order to determine whether an individual is identifiable the criteria provided by the Article 29 Working Party (hereafter, A29 WP) in its Opinion 05/2014 on anonymization techniques [22] are of interest. In this sense, the A29 WP introduces three criteria to analyze whether data can be considered as anonymous:

1. Singling out: refers to the possibility to isolate some or all records which identify an individual in the dataset.
2. Linkability: denotes the risk generated where, at least, two data sets contain information about the same data subject.
3.Inference: refers to the possibility to deduce, with significant probability, the value of an attribute from the values of other set of attributes.

In addition, it is essential to determine before which parties the data subject is identified or identifiable. In this part, we have to determine the parties involved in the service or data flow that identify or could potentially identify the individual as well as the specific risks that could cause the “discover” of data or the identification of the concrete individual. In the case of IdM services the analysis could be the following:

Table 1: Parties and risks of identification

Party	**Types of risks**
Third unauthorized parties	Information leaks, data theft or identity theft attacks
Service providers	Linkability risks, request of excessive personal data
Identity providers	Non-anonymized account or anonymized account but linkability between user’s attributes

Before each party, the specific techniques implemented in the technology subject of study must be considered. Depending on the existence or not of risks of identification, as well as their likelihood we could classify the "degree of anonymization" achieved before each party in the categories of low, medium or high. Nevertheless, note that only if a high degree of anonymization is achieved before all parties the DPIA could stop in this phase. Indeed, a medium degree of anonymization could be achieved in the case of making use of multiple pseudonyms but pursuing what established in Recital 26 of the GDPR, pseudonymized data shall be considered as personal data as it merely reduces linkability.

Determined that personal data are processed, the next step in the DPIA consists in the evaluation of risks. Risk management is necessary to determine the potential damages or risks to which an activity is exposed. From the perspective of data protection, the analysis focuses on those threats that affect rights and freedoms of individuals. As first step in the risk analysis we should classify threats depending on the risks source [23]. Our proposal would be to consider at least three risk sources:

1. Risks relating to the particularities of the service.
2. Risk relating to the architecture system components.
3. Risks relating to the user.

In the subject of study in this paper, IdM, these risks sources could materialize in the following threats:

Table 2: Risk sources and threats in IdM

Risk source	**Threats**
Risks relating to IdM services	Identity theft, information leaks, alteration of personal data.
Risks relating to the architecture system components	Malware diffusion, software and hardware failure, software manipulation, communication services failure, eavesdropping
Risks relating to users	Social engineering, extortion

As in the common methodology for the DPIA, the risk consists in the result of multiplying likelihood for impact. Likelihood and impact of threats are variable. We could make use of different scales for the likelihood and the impact that could be of quantitative or qualitative nature. The likelihood criteria could be classified regarding the frequency that the different threats might materialize (e.g., almost never, once per year, more than three times per year…). Regarding the impact criteria, the values could be associated with the deprivation of rights and freedoms.

Table 3: Likelihood and impact levels [Source: Spanish Data Protection Agency, 2018]

Likelihood		**Impact**	
Description	**Level**	**Description**	**Level**
Very likely	4	Maximum	7
Relevant	3	Significant	4
Limited	2	Limited	1
Unlikely	1	Negligible	0

Once determined the likelihood and impact for each specific threat, the results must be classified into different levels of risk. We could take as reference for the maximum level of risk the result of multiplying a very likely threat with a maximum impact. Conversely, the minimum level of risk would be obtained by multiplying an unlikely threat with a negligible impact. Between these two values we can create as many classifications as desired keeping a proportional relation between them.

Table 4: Risk levels [Source: PILAR,2020]

Risk levels	
Catastrophic	28
Disaster	25
Extremely critical	22
Very critical	19
Critical	16
Very high	13
High	10
Medium	7
Low	4
Negligible	0

Note that in the scenario of a technology that has not been implemented yet, there would exist values of impact that will not be possible to determine and will have to be restudied in a subsequent context-based analysis. Nevertheless, we propose for these cases to assign a medium level of impact (or in our scale, the level significant) that will refer to average data.

In addition, risks can involve different consequences or effects. In this sense, there might exist risks that affect the availability of the service, but that do not have privacy implications. In the study of privacy risks, three risk dimensions are of interest [24]:

1. Integrity of the data: data are correct and complete.
2. Confidentiality of the data: data remain unknown before unauthorized parties.
3. Authenticity of users and information: the user is the authorized person, and the information corresponds to this person.

Consequently, only those risks whose consequences affect these dimensions will be considered. The clearest example is the case of those risks affecting hardware components. Stealing physical hardware was a common method for committing cybercrime [25]. However, due to the existence of complex encryption processes this method lacks efficiency in some cases, affecting exclusively the availability of the service.

At this stage we must have already determined those risks with privacy implications that might affect our technological proposal. Depending on the features or characteristics of our technology, likelihood values must be assigned (e.g., in the case of the OLYMPUS technology due to the distribution of the task of the IdP and password fragmentation, identity theft was qualified as unlikely [26]). Concerning impact values, if we do not have knowledge about the specific data that will be processed (e.g., a technological proposal to be applied in the health sector), we recommend considering the values that will correspond to the processing of average data.

By performing this first assessment or in this first “layer”, we will obtain a set of results concerning the level of privacy by design achieved in a specific technological proposal. In the research project OLYMPUS, the level of privacy by design achieved by the technological proposal was studied by making use of the methodology proposed with certain modifications. In this study it was detected in an early stage that OLYMPUS suffered from two specific drawbacks in its conception or design [26]. On the one hand, its distributed architecture functioned by the replication of user’s attributes in each partial IdP which could increase information risks and challenge the principle of proportionality in the data processing. On the other hand, the authentication method supporting OLYMPUS solution was limited to text passwords, something that was problematic in the scenario of a highly resilient architecture which might favor attacks to focus on the user. The first problem is currently under study while the second one has already been solved by implementing a multifactor authentication process [27]. In addition, thanks to this preliminary study of privacy implications, technical and business partners were informed about the conditions of deployment to achieve the best privacy results.

This first assessment can conclude in three different ways:

a) The technical solution achieves an adequate level of privacy by design.
b) The technical solution still requires modifications or safeguards.
c) The technical solution will never achieve an adequate level of privacy by design.

If from this first study, safeguards are proposed, a deadline must be granted for the conception and implementation thereof. Note that in this phase it is extremely important a proactive and collaborative attitude between the members of the team. In this sense, legal experts must discuss the technical viability of the safeguards proposed with the experts in the area as well as the possible adaptation or modifications that these measures can suffer. Once the time for the implementation of safeguards has concluded, it should be reevaluated whether the technology has solved relevant risks. Consequently, we repeat the step performed before (i.e., the technical solution achieves an adequate level of privacy by design; or it still requires modifications; or it will never achieve an adequate level of privacy by design).

To conclude this subsection, it must be noted that privacy is not an absolute value and will have to be considered conjointly with the nature and scope of the data processed. Nevertheless, performing these prior assessments can bring important benefits as they do not only help to detect possible problems in the first stage of the development of the technology, but it can also advance whether a technology could or not be adequate for the processing of certain type of personal data.

3.2 Second layer

In the first and second scenario mentioned in the previous subsection (once safeguards have been introduced and reevaluated), the technology can now be implemented in a specific context. At this moment it is necessary to perform the second assessment or "layer" consisting in the study of the level of compliance with data protection principles in the specific scope of the data processing. Consequently, in this case the risks will refer to the lack of compliance with data protection principles. Likewise, in this second assessment we can also take into account other context-dependent aspects such as the training of the staff in charge of deploying the service. The development of this/these subsequent assessment/s will be easier since the implications of making use of a specific technology have already been determined. This second analysis is more similar to the traditional conception of the DPIA, as a methodology thought for legal professionals, and should include at least the following elements.

The first aspect to study is the lawfulness of the data processing. The data processing could be based on consent or any of the other circumstances envisaged in Articles 6 or 9 of the GDPR. In those cases where the data processing is based on user's consent, the data controller shall be able to demonstrate that user's consent was informed and given in a voluntary way. The same would apply to the procedure to withdraw his/her consent. On the other hand, for those cases where the processing will not be based on consent of the user, the data processing must be justified according to any of the other causes listed in the Article 6.1 of the GDPR and specific regulations must be taken into account (e.g. the Directive EU 2016/680 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties [28]).

Once the lawfulness of the data processing has been determined, other circumstances must be considered. In this sense, the data minimization principle implies to process the strictly necessary data for the justified purposes, as well that the storage will be limited to the strictly necessary term. In the risk analysis, examples of threats of lack of compliance with data minimization principle could be the excessive data collection or an excessive period of storage.

Likewise, proportionality in the data protection activity must be analyzed. From this perspective it is necessary to evaluate if the aim pursued with the data processing can be achieved by other means which imply a lower risk. In order to determine the proportionality of a data processing activity, the guidelines provided by the Spanish Data Protection Agency [23] might be useful. Pursuing these guidelines, three successive assessments are proposed:

1. If the measure can achieve the proposed objective (suitability criteria).

2. If no more moderate measure can meet this goal with the same effectiveness (necessity criteria).
3. If the measure implies more benefits than damages for other assets or values in conflict (proportionality criteria in strict sense).

The evaluation of the proportionality in a data processing activity does not only refer to a single threat, but it requires to considerate the set of elements characterizing the data processing. A good example in these cases could be the implementation of biometric authentication. Biometric data are qualified as sensitive data and they present a high risk since once compromised, they will be compromised forever [20]. However, biometric data also represent important advantages in the process of binding identity during authentication and it is extremely convenient for end users. Consequently, in the case of implementing a technology involving the use of biometrics in this “second layer” it must be studied whether the deployment of that specific technology is justified or balanced for the concrete scenario pursuing the criteria cited above.

In addition, accuracy, integrity, transparency and confidentiality in the data processing must be studied. Pursuing Article 5.1. of the GDPR, accuracy in the data processing requires data to be accurate in order to assure a correct fulfillment of requests and rights, as well as the possibility of the user to demand the correction of inaccurate data as stated in Article 16 of the same text [20]. On the other hand, transparency in the data processing can be considered from different perspectives. It means that the data subject can access his/her data at any moment with no need to provide special justification, but also that the data processing must be carried out in a way it enables and facilitates eventual controls by Law Enforcement Authorities. To conclude, confidentiality of the data implies that data must remain unknown before non-authorized parties, thus it requires appropriate mechanisms to assure that the person accessing the data is the authorized user.

Besides, in this second assessment we will count on additional information such as the staff in charge of providing the service or specific security measures adopted that could modify the initial result of our DPIA. By way of example, we will invent a use case where OLYMPUS technology could be deployed. We have noted the following aspects with regard to OLYMPUS:

a) It increased the amount of data process as it replicates user’s attributes in each partial IdP.
b) It was exclusively based on passwords.

As the second problem has been solved, we should have already changed the likelihood of social engineering attacks in our first DPIA.

The invented use case where we would deploy OLYMPUS is the following:

“OLYMPUS technology is implemented to provide services of identification and identity management (identity as a service) in the context of authentication before streaming services. In this case the data collected will be the name and surname of the user as well as his/her age and email address. User consent is obtained in a comprehensible and informed way. These data will be exclusively used for the purpose of providing identification services and will be erased in the moment the user decides to delete his account. The user can access his account and visualize his/her data at any moment. Financial information (i.e., credit card information) remains in the side of the service provider. The IdP does not receive/store any information about the content visualized”.

Table 5: Risk analysis in the "second layer"

Threat	Likelihood	Impact	Risk
Problems related to the lawfulness of data collection and processing	Unlikely	Significant	Low
Problems related to the transparency of the processing	Unlikely	Limited	Negligible
Problems related to excessive data collection	Unlikely	Limited	Negligible
Problems related to accuracy of the data	Unlikely	Limited	Negligible
Problems related to the retention period of the data	Unlikely	Limited	Negligible
Problems related to the rights of the interested subject	Unlikely	Significant	Low
Unauthorized access to personal data	Relevant	Limited	Negligible
Impersonation of the user	Unlikely	Maximum	Medium
Profiling	Limited	Significant	Medium

Considering that financial information remains in the side of the service provider (the streaming service), the nature of the data processed by the IdP make risk of unauthorized access limited. Conversely, impersonation of the user will allow access to financial information and the content visualized, hence in case of materialization of this threat the impact would be maximum. Nevertheless, the resulting risk of impersonation must be considered as medium thanks to OLYMPUS distributed architecture that reduces the likelihood of this risk in common IdPs.

This analysis can be repeated in different use cases where the OLYMPUS technological proposal aims to be implemented. The process will be easy as the previous analysis performed with regard to the technological proposal for IdM has already defined the likelihood of those risks commonly linked with IdM services for this specific technology.

4 CONCLUSIONS

The methodology exposed along this paper evidences the need of adopting multidisciplinary approaches that involve the collaboration of experts from different areas in the development of safer, more privacy-respectful and human-centered technologies. More specifically, we have proposed a multiphase DPIA, or in other words, the division of the DPIA methodology in two phases or "layers" to obtain more efficient results and avoid problems such as the ones described in the introduction and along this paper (i.e., vulnerable technologies, wide rejection of technological proposals or implementation of technological proposals involving a high privacy risk). Indeed, a multiphase DPIA would be more adapted to the evolutive reality of a technological project. In this sense, it is not enough to determine a set of requirements and definitions at the beginning of a project, but legal compliance must be dynamic and adapted accordingly with the technology's evolution.

We consider that this new approach is necessary because the DPIA, despite being conceived as a legal tool in its design, involves some technical aspects difficult to understand for legal experts. Conversely, it also involves legal concepts that technical experts are not used to manage. In consequence, we propose an innovative and interdisciplinary approach that favors the collaboration between both professionals and assures privacy not only since the earliest stage in the development of the technology, but also during the technology evolution and adaptation.

Nevertheless, it must be noted that the DPIA is not a tool designed to obtain absolute values but to study the level of balance achieved. This is particularly important when we are considering the evaluation of a technology that has not been implemented yet. Therefore, in the development of the first "layer", conclusions have to be interpreted and compared with other technologies, the current state of the technique, or complemented with subsequent analysis of real use cases or examples of second layers.

In conclusion, technological developments cannot ignore social or economic realities, as well as the protection of rights and freedoms established by regulations. Therefore, although perfect solutions do not exist, all these aspects must be studied in order to support the society to evolve in the adoption of technologies that are safe and respectful with the rights of the individuals but also with the society as a whole. Digitalized societies have come to stay. However, the future of technology is multidisciplinary and legal compliance must adapt to these challenges so that Law is not perceived as a barrier for innovation but as an essential safeguard for the protection of fundamental rights and civil liberties.

ACKNOWLEDGMENTS

[*To complete after*]

REFERENCES

[1] In Cambridge Analytical scandal it was discovered that Facebook provided unauthorized access to personally identifiable information of more than 87 million Facebook users to the data firm Cambridge Analytical. Cambridge Analytica integrated this information with a range of data from social media platforms, browsers, online purchases, voting results, and more. By adding OCEAN analysis to the other private and public data acquired, Cambridge Analytica developed the ability to "micro-target" individual consumers or voters with messages most likely to influence their behavior. The OCEAN analysis was paired with a large number of targeted messages in "Project Alamo," which was employed for the election campaign of President Trump. Jim Isaak and Mina.J Hanna. 2018. User Data Privacy: Facebook, Cambridge Analytica, and Privacy Protection", Computer 51(8), 56–59. 10.1109/MC.2018.3191268

[2] Jon R. Knight. 2019. The New Normal: Easier Data Breach Standing Is Here to Stay" *Cybersecurity L.Rep* 1 Feb. 6. Retrieved the 25th 2021 of April from: https://perma.cc/QXZ8-JEH3

[3] Paige Leskin. 2018. The 21 Scariest Data Breaches of 2018.*Bus. Insider* Dec. 30. Retrieved the 25th of April 2021 from: https://www.businessinsider.com/data-hacks-breaches-biggest-of-2018-2018-12

[4] Penny Jorna, Russel Smith and Katherine Norman. 2018. Identity crime and misuse in Australia: results of the 2018 online survey [Online statistical report]. Australian Institute of Criminology, 2020. Retrieved the 4th of April 2020 from: https://aic.gov.au/publications/sr/sr19

[5] The United States Department of Justice. Identity theft. Official website. Available at the following address: https://www.justice.gov/criminal-fraud/identity- theft/identity-theft-and-identity-fraud.

[6] Sopna A/P Sinnathamby Sehgar and Zuriati Ahmad Zukarnain, 2021. Online Identity Theft, Security Issues, and Reputational Dam- age. *Preprints* 1–10, 2021. https://doi.org/10.20944/preprints202102.0082.v1

[7] Tiffany Hsu, 2017, Data Breach Victims Talk of Initial Terror, Then Vigilance. N.Y. *TIMES* Sept. 9. Retrieved the 25th of April 2021 from: https://www.nytimes.com/2017/09/09/business/equifax-data-breach-identity-theft-victims.html

[8] Susan Gasson. 2003. Human-centered vs. user-centered approaches to information system design, *JITTA* 29-46. Retrieved the 26th of April 2021 from: https://aisel.aisnet.org/jitta/vol5/iss2/5/

[9] https://olympus-project.eu/(Grant Agreement 786725)

[10] Jorge Bernal, Antonio Skarmeta, Rafael Torres et al.. 2019. D3.1- Requirements and Design Templates for OLYMPUS. *Horizon 2020 Project OLYMPUS (Oblivious identitY Management for Private and User-friendly Services)*. Retrieved the 26th of April 2021 from:

https://olympus-project.eu/wp-content/uploads/2019/07/Olympus_pu_d3_1_v1.0.pdf

[11] Nabie Y. Conteh1 and Paul J. Schmick. 2016. Cybersecurity: risks, vulnerabilities and countermeasures to prevent social engineering attacks. *International Journal of Advanced Computer Research* 6(23), 31–38. https://doi.org/10.19101/ijacr.2016.623006

[12] Hitoshi Kokumai. 2018. Identity Assurance by Our Own Volition and Memory Part 1. *Payments Journal* 1st August. Retrieved the 26th of January 2021 from: https://www.paymentsjournal.com/identity-assurance-by-our-own-volition-and-memory-part-1

[13] Hitoshi Kokumai, 2019. Passwords Made of Unforgettable Images. *Payments Journal* 30th September. Retrieved the 26th of January 2021 from: https://www.paymentsjournal.com/passwords-made-of-unforgettable-images/

[14] Hitoshi Kokumai, 2021. Detection of Phishing by Episodic Image Memory. *Hitoshi Kokumai LinkedIn profile*

[15] Wan Ying Lee. Chee-Seng Tan and Poh Chua Siah. 2017. The Role of Online Privacy Concern as a Mediator between Internet Self-Efficacy and Online Technical Protection Privacy Behavior", *Sains Humanika* Vol.*9* no.3-2 37-43. Retrieved the 26th of April 2021 from: https://sainshumanika.utm.my/index.php/sainshumanika/article/viewFile/1271/724

[16] Premier Ministre. 2012. Note technique Recommandations de sécurité relatives aux mots de passe [Technical note]. *ANSSI*. Retrieved the 26th of April 2021: https://www.ssi.gouv.fr/uploads/IMG/pdf/NP_MDP_NoteTech.pdf

[17] Argyri Pattakou, Aikaterini-Georgia Mavrodei, Vasiliki Diamantopoulou et al.. 2018. Towards the design of usable privacy by design methodologies. In *Proceedings - 2018 5th International Workshop on Evolving Security and Privacy Requirements Engineering,* ESPRE 1–8. https://doi.org/10.1109/ESPRE.2018.00007

[18] Communication from the Commission to the European parliament and the Council on Promoting Data Protection by Privacy Enhancing Technologies (PETs) Brussels, 2.5.2007 COM (2007) 228 final. Available online: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52007DC0228&from=EN

[19] Spanish Data Protection Agency. 2019. A Guide to Privacy by Design. Issue October 2019.Retrived the 27th of April from: https://www.aepd.es/sites/default/files/2019-12/guia-privacidad-desde-diseno_en.pdf

[20] Regulation (EU) 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 the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) Official Journal of the European Union, Vol.L119 (4th May 2016). Available online: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679&from=EN

[21] Michèle Finck and Frank Pallas. 2019. They who must not be identified- Distinguishing Personal from Non-Personal Data under the GDPR. *Max Planck Institute for Innovation and Competition Research Paper Series* No.19- 14, .2-21. http://dx.doi.org/10.2139/ssrn.3462948

[22] A29 WP, 2014: "Opinion 05/2014 on Anonymization Techniques." Adopted on 10th April. WP 216. 0829/14/EN. Available online: https://cnpd.public.lu/dam-assets/fr/publications/groupe-art29/wp216_en.pdf

[23] Spanish Data Protection Agency. 2018. Guía práctica para las evaluaciones de Impacto en la protección de los datos sujetas al RGPD (25th May 2018). Retrieved the 27th of April 2021 from: https://www.aepd.es/sites/default/files/2019-09/guia-evaluaciones-de-impacto-rgpd.pdf

[24] These risk dimensions sources have been extracted and adapted from the information management tool PILAR. PILAR is a tool for information security management. More information available at National Cryptologic Centre website: https://www.ccn-cert.cni.es/pdf/guias/series-ccn-stic/400-guias-generales/2133-ccn-stic-470-h1-manual-de-la-herramienta-de-analisis-de-riesgos-pilar-6-2/file.html

[25] Judge Mohamed Chawki and Dr. Mohamed S. Abdel Wahab. 2006. Identity theft in cyberspace: Issues and solutions. *Lex Electronica,* vo.11 no.1.Retrived the 27th of April 2021 from: https://papyrus.bib.umontreal.ca/xmlui/bitstream/handle/1866/9563/articles_54.pdf?sequence=1

[26] Ignacio Alamillo, Cristina Timón and Julián Valero et al. 2020. D3.2- "Security and Privacy-aware OLYMPUS Framework Impact Assessment". Horizon 2020 Project OLYMPUS (Oblivious identitY Management for Private and User-friendly Services, 2020 [Online deliverable] Retrieved the 27th of April 2021: https://olympus-project.eu/wp-content/uploads/2020/02/Olympus_pu_d3_2_v1_0.pdf

[27] Anja Lehmann, Rafael Torres et al.2020. D3.3 OLYMPUS Blueprint. *Horizon 2020 Project OLYMPUS (Oblivious identitY Management for Private and User-friendly Services*, 2020 [Online deliverable] Retrieved the 27th of April 2021: https://olympus-project.eu/wp-content/uploads/2020/10/Olympus_pu_d3_3_v1_0.pdf

[28] Directive EU 2016/680 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties. Official Journal of the European Union, Vol.119/89 (27th April 2016). Available online: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016L0680&from=EN