# AI Training Datasets & Article 14 GDPR: A Risk Assessment for the Proportionality Exemption of the Obligation to Provide Information

*Submitted*: 29 January 2021
*Revised*: 3 March 2021
*Accepted*: 1 April 2021

Iakovina Kindylidi[*]
ORCID: https://orcid.org/0000-0002-8803-1359
Inês Antas de Barros[**]
ORCID: https://orcid.org/0000-0002-1226-1289

*Article submitted to peer blind review*

DOI: https://doi.org/10.26512/lstr.v13i2.36253

**Abstract**

**[Purpose]** At the earliest stages in AI lifecycle, training, verification and validation of machine learning and deep learning algorithm require vast datasets that usually contain personal data, which however is not obtained directly from the data subject, while very often the controller is not in a position to identify the data subjects or such identification may result to disproportionate effort. This situation raises the question on how the controller can comply with its obligation to provide information for the processing to the data subjects, especially when proving the information notice is impossible or requires a disproportionate effort. There is little to no guidance on the matter. The purpose of this paper is to address this gap by designing a clear risk-assessment methodology that can be followed by controllers when providing information to the data subjects is impossible or requires a disproportionate effort.
**[Methodology]** After examining the scope of the transparency principle, Article 14 and its proportionality exemption in the training and verification stage of machine learning and deep learning algorithms following a doctrinal analysis, we assess whether already existing tools and methodologies can be adapted to accommodate the GDPR requirement of carrying a balancing test, in conjunction with, or independently of a DPIA.
**[Findings]** Based on an interdisciplinary analysis, comprising theoretical and descriptive material from a legal and technological point of view, we propose a risk-assessment methodology as well as a series of risk-mitigating measures to ensure the protection of the data subject's rights and legitimate interests while fostering the uptake of the technology.
**[Practical Implications]** The proposed balancing exercise and additional measures are designed to facilitate entities training or developing AI, especially SMEs, within and

[*] Iakovina Kindylidi is an international adviser at Vieira de Almeida & Associados' ICT practice area. She holds an LL.M in International Business Law from Tilburg University and has participated as speaker in various seminars and classes on emerging technologies, with a focus on AI. E-mail: imk@vda.pt. Address: Rua Dom Luís I 28, 1200-151 Lisboa.
[**] Inês Antas de Barros is a managing associate at Vieira de Almeida & Associados' ICT practice area. She holds an LL.M in International Business Law from Global School of Law of the Catholic University of Portugal. She has participated as a speaker in various seminars and classes on privacy, data protection, and cybersecurity. E-mail: iab@vda.pt.

KINDYLIDI, I.; BARROS, I. A. de. *AI Training Datasets & Article 14 GDPR: A Risk Assessment for the Proportionality Exemption of the Obligation to Provide Information*. **The Law, State and Telecommunications Review**, v. 13, no. 2, p. 1-27, October 2021.

outside of the EEA, that wish to ensure and showcase the data protection compliance of their AI-based solutions.

**Keywords**: AI. GDPR. Article 14. Risk-Assessment. Transparency.

## INTRODUCTION

Artificial Intelligence ("AI") and Big Data are topics of high priority for the European Commission as it recently published its holistic regulatory proposal on AI[1], while further promoting data sharing in the EU with the creation of the European data spaces in key sectors[2]. As algorithms and AI-related technologies are fueled by data, one of the fundamental concerns of academics and regulators has been the data protection of individuals.

Although the data protection and privacy risks stemming from the use of AI are manifold and undoubtfully crucial, legal scholars are focusing more on how these risks can be avoided - preventively or punitively – than on how can entities proactively build GDPR-compliant AI.

At an earlier stage in the AI lifecycle large datasets which usually include personal data are "fed" to machine learning[3] and deep learning[4] algorithms to support their training and functions as well as to test the AI's behavior in the subsequent stages of verification and validation. This personal data is usually not obtained directly by the data subjects but via third parties, private or publicly available sources, and even if the data subject is identifiable, the process to identify them may be too difficult and, most of the time, of no interest to the entity training the algorithm[5]. Nonetheless, this data, even if it may not directly identify a data subject, when related to other datasets may lead to the identification of a person.

Furthermore, in relation to the algorithm training stage, although due to several incidents there is conversation around AI-biases and the importance of data quality to mitigate such errors and fostering fundamental rights, there is little discussion on how, in practical terms, can an entity use these datasets to train and

[1] Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) And Amending Certain Union Legislative Acts COM/2021/206.

[2] Proposal for a Regulation Of The European Parliament And Of The Council on European data governance (Data Governance Act) COM/2020/767.

[3] Machine learning, is the ability of an algorithm to improve its performance autonomously, based on newly acquired information and experience.

[4] Deep learning is a subset of machine learning. Deep learning enables the algorithm to make decisions through data processing and creation of patterns.

[5] AEPD is referring to this data as "quasi-identifiers", also mentioned as pseudo-identifiers or indirect identifies. See amongst others (AEPD, 2021, p. 33).

run its algorithms pursuant to the GDPR principles and respecting the rights of the data subjects. Recently, however, it was proposed that processing of data in the training stage should be distinguished from the operational part (SARTOR e LAGIOLA, 2020, p. 80).

In the light of the above, our analysis will focus only on the stage of training and developing the AI and on the stage of verification and validation. We will address the obligation of the controller to provide information to the data subjects pursuant to the transparency principle when the data is not directly collected by the data subject and attempt to answer how can an entity using large datasets containing personal data for the purpose of AI training comply with such obligation. Additionally, we will assess possible solutions for those entities that reasonably cannot meet such obligation but still wish to be GDPR-compliant following a risk-assessment analysis. Lastly, we will propose a balancing methodology and a set of additional technical and organisational measures that can be used by organisations that wish to mitigate their data protection risks when training and developing AI.

For clarity it should be noted that in the present when we refer to Artificial Intelligence ("AI") we refer only to machine learning or deep learning AI and the terms AI or algorithm are used interchangeably. All the articles referred herein are articles of the General Data Protection Regulation (EU) 2016/679 ("GDPR").

## ARTICLE 14 AND THE OBLIGATION TO PROVIDE INFOMRATION

### Introduction

The transparency principle is embedded in Articles 12, 13, 14 and 34 GDPR. The transparency principle mandates a higher level of transparency by the data controller regarding their processing activities, including the collection, use and sharing of data. In accordance with Recital 39, the information provided to the data subjects should be sufficient and permit the individuals to effectively exercise their rights under Articles 15 to 22 GDPR.

The principle of transparency is intertwined with the principles of lawfulness and fairness. The obligation to provide information to the data subjects appears as a condition for fair and lawful processing while at the same time ensures compliance with the transparency principle (ZANFIR-FORTUNA, 2020, p. 415).

The importance of transparency is paramount when complicated emerging technologies are carrying the processing of data, such as AI. The opacity or blackbox[6] problems of machine learning and deep learning algorithms create the need for a concrete data governance strategy that strikes a balance between GDPR compliance and the characteristics of machine learning algorithms (KINDYLIDI, 2020, p.121-123; KROLL, HUEY, et al., 2017). At the same time, aside from an ex ante obligation to provide information regarding the automated decision-making, a vivid academic debate ensued on whether there is an obligation of the controller to provide an ex post explanation for the processing involving AI (CABRAL, 2021; MAGLIERI e COMANÉ, 2017; WACHTER, MITTELSTADT e FLORIDI, 2017).

Applicable to any processing activity, including automated-decision making, Article 12 further outlines the scope of the transparency principle and, concomitantly, the obligation of the data controller to provide information to data subjects. Article 12 does not specify the manner and form that can be used – except when specifically addressed to children. However, it obliges the controller to provide information, in every communication with an individual regarding personal data (Article 5.1 lit. a GDPR), in a concise, transparent, intelligible and easily accessible form, using plain language about its operations involving personal data[7].

The principles of fairness and transparency require that the information provided to the individual include the existence of any processing operation on their personal data as well as the legal basis and purposes of such processing (Recital 60 GDPR). In other words, minimum information on the processing should be provided, irrespective of whether the controller directly collects the data by the data subjects or they are obtaining it via a third party, although exceptions to the general rule of providing information at the time of collection apply, as well as on the specific information to be provided to the data subject.

Although the communication of information is not subject to specific strict formal requirements, information should be provided upon collection or when the personal data is obtained by the controller using appropriate means, including electronic means especially when the personal data is processed by electronic means or obtained online. In practical terms, electronic means is generally preferable in most processing activities and it is particularly relevant when it comes to AI training datasets. Ipso facto, it is likely that an entity processing

[6] The deep neural networks with millions of connections which all combined form the decision of the AI, together with the statistical based truths on which the algorithms run, make it particularly difficult to trace back the decision and verify whether it functions properly, thus, detecting, and concomitantly, correcting possible errors of the system can be a difficult task (KINDYLIDI, 2020, p. 121).

[7] For instance, as mentioned in Recital 59, one of the means that the controller can deploy is to allow electronic requests of individuals regarding their data, especially when the data is processed also by electronic means. the controller should provide means for.

personal data for the training of an algorithm will use electronic means to provide information to the data subject[8].

Following the general requirements of Article 12, Article 13 obliges the controller to proactively inform the data subjects about the processing activity when the controller is the one directly collecting the data and at the time of such collection. The information is usually contained, for instance, in a notice, statement or the privacy policy on the website of the controller.

Although our analysis is focusing on Article 14, hence, the obligation provides information when the data is not directly collected by the data subjects, the guidance already provided by the ECJ case law and the work of the European Data Protection Board ("EDPB") on Articles 12 and 13 can be of assistance on the interpretation of the principle of transparency and the common aspects of Articles 13 and 14.

## Article 14: An Overview

Article 14 regulates the obligation to provide information when the data is not directly collected by the data subjects. As a reflection of the principle of transparency, Article 14 requires from the data controller to proactively disclose information regarding the processing of data.

In addition to the information required under Article 13, the controller is required to also specify the categories of personal data obtained (e.g. biographical, behavioral, financial, sensitive etc.) as well as its source, including publicly available sources, unless it is not possible (WP29, 2017). Notwithstanding, according to the transparency guidelines of Article 29 Working Party ("WP29"), even if it is impossible to specify the exact source, for example due to confidentiality obligations or because the source is not known to the controller, at least a general reference to the characteristics of the source, i.e. whether it is publicly or privately held or of the type of the third party source, i.e. organisation, industry or sector, should be included (WP29, 2017).

Moreover, a list of the information that should be provided to the data subjects is included in paragraphs 1 and 2 of Article 14. To a large extent the information is the same as this of Article 13. More specifically, in addition to the information of Article 13, considering the specific circumstances of Article 14, when the data is not collected directly from the controller, the controller should also provide information about the categories of personal data obtained (Article 14.1.lit.d) and its source (Article 14.2.lit.f). The table below offers an overview of the type of necessary information and when they should be provided pursuant to Articles 13 and 14:

[8] See below paragraph *Article 14: An Overview*.

| **Type of Information to Be Provided** | **Article 13** | **Article 14** |
|---|---|---|
| Name and Contact Details of Controllers | Always | Always |
| Name and Contact Details of Representative | If Applicable | If Applicable |
| Name and Contact Details of DPO | If Applicable | If Applicable |
| Purpose of Processing | Always | Always |
| Lawful Basis for Processing | Always | Always |
| Legitimate Interests for Processing | If Applicable | If Applicable |
| Categories of Personal Data Obtained | **Not Required** | Always |
| Recipients or Categories of Recipients of Data (*e.g.* Processors) | If Applicable | If Applicable |
| Transfer of Data to Third Countries | If Applicable | If Applicable |
| Retention Period | Always | Always |
| Right Available to Data Subjects | Always | Always |
| Right to Withdraw Consent | If Applicable | If Applicable |
| Right to Lodge a Complaint with a Supervisory Authority | Always | Always |
| Source of the Personal Data | **Not Required** | Always |
| Information Regarding Automated Decision-Making Proofing | If Applicable | If Applicable |

**Table 1 –** Type of Information to Be Provided under Articles 13 and 14 GDPR

Regarding automated decision making, and irrespective of whether automated decision-making falls within the definition of Article 22.1 or not, it is suggested by WP29, as a good practice, that the following information should be provided when automated decision-making is deployed to ensure fair processing (WP29, 2018, p. 24-25):

- Inform individuals about engaging in automated-decision making or profiling;
- Provide meaningful information about the logic involved; and
- Explain the significance and foreseen consequences of such processing.

All this information should be provided to the data subjects within a reasonable period and not later than one (1) month from obtaining the data. Within this framework, the controller is free to decide the specific time for sending the information notice, while taking into consideration the specific circumstances of the processing. In this regard, there are two specific situations that have a shorter deadline:

- When the data will be used for communication with the data subject, the information should be provided at the time of the first communication; and
- If the data will be disclosed to another recipient, the information should be provided when the data is first disclosed.

Although, these specific circumstances need to be assessed on a case-by-case basis, usually, they will not be met in the scenario of AI training or validation/verification.

Moreover, four exemptions to the obligation to provide information are identified in Article 14.5:

- When the data subject has already been informed regarding this processing, as for instance in situations where at the time of collection of data the data subject was also informed about this further processing, in this case the processing for the purposes of training AI;
- When it is impossible or it requires disproportionate effort for the controller to provide the information to the data subjects, in particular for processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes (Article 89), under certain circumstances;
- Where disclosure or obtaining data is expressly laid down by Union or Member State law, provided that the law offers adequate measures to protect the legitimate interests of the data subjects; and
- Where the personal data must remain confidential, subject to professional secrecy obligation regulated by Union or Member State Law as for example in the case of lawyers or medical practitioners.

In the following paragraph we will focus our analysis on the fourth exemption introduced in Article 14.5. lit. b but when the data processing is not carried for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes.

## Proportionality Exemption to the Obligation to Provide Information

### *Overview*

As stated above, under Article 14 the controller should proactively provide information to the data subjects. This means, that the controller must seek to inform the data subjects about the processing, as opposed to simply making the information publicly available together with the terms and conditions of the controller's website (WP29, 2017, p. 16-17). However, this may not be feasible when there is no direct link between the controller and the individual and when the data available may not allow an easy identification of the data subject. This will be the case, in particular, for processing large datasets for AI training or verification and validation.

In Article 14.5. lit. b there is an exemption introduced for when it is impossible to deliver information or when the delivery of the notice will require disproportionate effort. The controller has the burden of proof for demonstrating that impossibility (WP29, 2017, p. 28-29).[9]

Although this exemption applies to any type of processing, WP29 has stated that the exemption of disproportionate effort should exceptionally be used by controller "who are not processing personal data for the purposes of archiving in the public interest, for scientific or historical research purposes or statistical purposes" (WP29, 2017, p. 30). For what is more, according to WP29 all the exemptions to the obligation to provide information should be interpreted narrowly (WP29, 2017, p. 28). Therefore, even if a controller falls within the scope of one of those exemptions it should do so following a thorough risk-based analysis.

In this regard, WP29 suggests that in the case of disproportionate effort a balancing exercise should be undertaken between the effort required by the controller to notify the data subjects and the possible risks for the data subjects if such information is not provided while this assessment should be thoroughly documented by the controller as part of their accountability obligations (WP29, 2017, p. 31). The specific process for such assessment, however, is not specified.

Furthermore, the jurisprudence of European Court of Justice ("ECJ") cannot provide guidance on the exemption of the information obligation since the available case law regarding the transparency principle and the obligation to inform the data subjects does not refer specifically to indirect collection of personal data (C-201/14, Bara) or any reference is incidental (C-212/13 Ryneš). For example, in Google Spain the ECJ stated that it is likely that, when in doubt, the rights and interests of the data subject in receiving information will overrule

[9] For the opinion that this exemption should be extended also when the data were collected directly by the data subject see (SARTOR e LAGIOLA, 2020, p. 54).

the controller's interest in not providing the information (C-131/12, Google Spain).

Similarly, aside from the transparency guidelines of WP29 (WP29, 2017), neither EDPB nor any national supervisory authorities have issued any specific guidance on Article 14 and the exemption of paragraph 5.lit.b. Notwithstanding, it should be noted that the Information Commissioner's Office ("ICO") guide to GDPR contains two examples for the exercise of the right to provide information in practice, which can provide further guidance:

- "Data obtained from publicly accessible sources: Controller needs to provide the individuals in any case with privacy information, especially in relation to any unexpected or intrusive use of their data (e.g. combining information) within a reasonable period from obtaining the data and not later than 1 month, unless if an exemption of Article 14.5 GDPR applies. If the controller thinks that it is impossible to provide privacy information to the individual or it would involve a disproportionate effort, they must carry out a data protection impact assessment ("DPIA") to identify measures to mitigate the risks for the data subjects.
- Buying personal data from other organisations: Controller is required to provide its own privacy notice within a reasonable period from obtaining the data and no later than 1 month, except if they fall within the exemptions of Article 14.5 GDPR. Note that if the purpose for using the personal data is different to that for which it was originally obtained controller should notify individuals and provide and disclose the lawful basis for the processing. As in (a), a DPIA must be carried out if the controller thinks that it is impossible to provide privacy information to the individual or it would involve a disproportionate effort".[10]

### *Article 14 & Article 11: Processing without Requiring Identification*

AI training requires vast datasets. Thus, it is usual, considering the size of information included and the various data sources that can be deployed, that the controller processing the data and training the AI is not aware of all the data included in the datasets, whether and to what extent there is personal data involved and most of the times and depending, of course, on the specific objective of the AI training[11], may not have any interest in or may not be in a position to identify

[10] (ICO, 2018, p. 98).

[11] For instance, an AI that will be used for profiling, by design is trained on various personal data belonging to the same data subject while an AI that will be used for identifying bird species it might be "fed" some photographs of humans to teach it to distinguish humans from birds but probably no additional personal data on the data subject will be provided.

the data subjects or such identification may result to disproportionate effort. In this regard, Article 11 GDPR may apply. Therefore, it is important to briefly refer to the interplay between Articles 11 and 14.

More specifically, under Article 11, if the purpose for which a controller processes personal data does not require (or it does not require any longer) the identification of the data subject, then the controller is not obliged to maintain, acquire or process any additional information to identify the data subject for the sole purpose of complying with GDPR, for instance with Article 14, in order to allow data subjects to exercise their rights. In other words, Article 11 applies when the controller holds personal data, but some informational elements are missing, and the controller cannot identify the data subject.

Pursuant to the core principles of purpose limitation, data minimisation and storage limitation, Article 11 grants the advantage to the data controller that has implemented the appropriate technical and organisational measures, such as pseudonymisation techniques (Articles 25 and 32), the controller is not required to provide any additional information to the data subjects in cases where identification is not or no longer possible as for example in the case of AI training datasets.

Moreover, an exemption from the rights of the data subjects (Articles 15-20) is granted if the controller is able to demonstrate (if requested ex post by a supervisory authority or court) that it is unable to identify the data subject (GEORGIEVA, 2020, p. 396). “This significant exemption” (FRA, 2018, p. 94) is however limited. If the data subject reaches out to the controller in order to exercise their rights the controller cannot refuse based on the exemption of Article 11 to take additional information provided by the data subject in order to support the exercise of their rights. Nonetheless, without prejudice to the operational objective of the AI, the scenario of the individual whose data is included in the datasets used for the AI training or validation contacting the controller is not very probable.

Due to the uniqueness of Article 11 GDPR, as it is an original provision of the GDPR, currently, there is no relevant case law. Nonetheless, the ECJ jurisprudence on the principles of purpose limitation, data minimisation and storage limitation are of relevance (C-131/12, Google Spain, para. 93; Case C-553/07, Rijkeboer, paras 67 and 70). In particular, Article 11 is in line with ECJ decision on Breyer, where it was suggested that it is not necessary that the controller alone holds all the information in order for the data to be qualified as personal data (C-582/14, Breyer, para. 43).

## BALANCING EXERCISE & ADDITIONAL MEASURES

### Risk-Assessment under Article 14.5.lit.b

In the previous paragraph it was stated that when the controller finds impossible to deliver the information to the data subjects or when the provision of the notice will require disproportionate effort, they should undertake a balancing exercise. The specific process for such an assessment is not, however, specified.

At this point, it should be noted, that of course the risk assessment will depend on the use and possible application of the AI once its trained. For instance, it may be possible that the AI training datasets may fall within the scope of scientific or historical research processing purposes or statistical purposes. Notwithstanding, as it is mentioned above, the exemption is rather narrowly defined.

Recital 62, although is referring to the example of processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes, can provide further guidance on the elements to be taken into consideration as part of the risk-assessment exercise (WP29, 2017, p. 30):

- The effort of the controller to notify the data subjects;
- The number of data subjects involved;
- The age of the data; and
- The safeguards that the controller has in place.

Concomitantly, pursuant Recitals 75 and 76, any risk analysis for the rights and freedoms of individuals should be based on an objective assessment that considers two main elements: the likelihood and severity of the risk to the rights and freedoms of data subjects. These two elements should be determined by a reference to the nature, scope, context and purposes of the processing at stake.

Furthermore, the only guidance coming from a national data protection authority is from ICO where in the examples included in its guide and referred above, a DPIA is recommended if the controller thinks that it is impossible to provide privacy information to the individual or it would involve a disproportionate effort to identify measures to mitigate the risks for the data subjects.

Under Article 35, DPIAs are mandatory for data processing operations likely to result in a high risk for the rights and freedoms of natural persons. A set of indicative examples if further provided from which two may be relevant for the processing activities undertaken for the training or validation and verification process of AI:

- When there is a systemic and extensive evaluation of personal aspects relating to natural persons which is based on automated processing, including profiling, and on which decisions are based that produce legal effects concerning the natural person or similarly significantly affect the natural person;
- Processing on a large scale of special categories of data referred to in Article 9(1) GDPR, or of personal data relating to criminal convictions and offences referred to in Article 10 GDPR.

Whether the processing activities undertaken in the training or validation and verification stages of AI lifecycle will depend of course on the specific datasets as well as the operational objectives of the AI. Nonetheless, it is arguable, however, if the processing activities of these stages seen alone fall within the indicative examples of Article 35. Furthermore, to assess whether a DPIA is mandatory or recommended in these situations it should be assessed whether not sending an information notice to the data subjects will amount to a high risk of for the rights and freedoms of natural persons.

WP29 in its DPIA Guidelines has further elaborated the GDPR-based characteristics of high-risk processing operators and has suggested a self-assessment checklist of nine (9) criteria (WP29, 2017, p. 9-10). If two (2) out of these nine (9) criteria are met, then it is likely that a DPIA is necessary while the more criteria are met, the more likely it is for the processing activities to present a high-risk for the rights of the data subjects. The table below presents an overview of the WP29 criteria and a short assessment in the context of AI training and validation processing.

| **WP29 Criteria** | **Assessment** |
|---|---|
| Processing involving evaluation or scoring, including profiling and prediction the data subject's performance (*e.g.* at work, economic situation, health etc.) or interests or behaviour (*Recitals 71 and 91*) | Depending on the objective of the AI and the amount of data involving one person used it may apply.<br>Training or validation processing activities as such do not meet this criterion. |
| Processing involving automated-decision making with legal or similar significant effect (*Article 35.3.lit.a*) | Depending on the objective of the AI.<br>Training or validation processing activities as such seem to have little or no effect on individuals without considering the operational stage of AI. (*WP29, 2018, p. 11-12 and 21-22*) |

| | |
|---|---|
| Processing involving systematic monitoring of data subjects (*Article 35.3.lit.c.*) | Depending on the objective of AI and the means of collecting data it may apply. |
| Processing involving sensitive data or data of highly personal nature (*e.g.* information on political opinions, criminal convictions or offences etc.) (*Articles 9 and 10*) | Depending on the nature of data included in the datasets together with factors such as whether this data was made publicly available by the individual it may apply. |
| Processing on a large scale based on factors such the number of data subjects concerned, volume of data, duration of processing, geographical extent of the processing (*Recital 91*) (*WP29, 2016, p. 7-8*) | By design, it will usually apply in AI training and validation processing activities. |
| Processing involving matching or combining different datasets perhaps coming from different controllers (*WP29, 2013, p. 24*) | By design it will usually apply in AI training and validation processing activities although the objective of the AI may affect this assumption. |
| Processing of data concerning vulnerable data subjects (*e.g.* children, elderly, patients etc.) (*Recital 75*) | Depending on the objective of the AI and the type of data included in the datasets it may apply. |
| Processing involving innovative use or applying new technological or organisational solution that may trigger novel forms that may have a direct impact on individual's privacy (*Article 35.1 and Recitals 89 and 91 GDPR*) | By design it will usually apply in the AI training and validation processing activities. |
| Processing in itself preventing data subjects from exercising a right or using a service or a contract (*Article 22 and Recital 91*) | Depending on the objective of the AI it may apply. |
| **Final Score** | **39** |

**Table 2 –** Number of people treated per DPIA Criteria and Assessment

As it is illustrated in the table above, training and validation processes, as such, without taking into consideration the specificities and objectives of each processing0, fulfil three (3) out of the nine (9) criteria set by WP29. This assessment translates to a possible – not mandatory – need for a DPIA. National supervisory authorities may approve white lists for mandatory DPIAs, under which training, and validation processes may always require a DPIA. Nonetheless, the assessment carried out following WP29 criteria does not translate necessarily to a high-risk activity for the data subjects. In the case when the controller does not consider that the processing is "likely to result in a high risk" for the individuals, the controller should duly justify and document the reasons for not carrying out a DPIA, and where applicable record the opinion of their data protection officer ("DPO") (WP29, 2017, p. 12)

In the light of the above, irrespective of whether carrying a DPIA in these situations is determined as an obligation of the controller or as a recommendation, the WP29 DPIA guidelines can support the risk-assessment exercise of the controllers. At the same time carrying a DPIA can assist the controller in assessing the risks for the data subjects arising not only from the processing activities but also, in particular, from not providing an information notice to the data subjects. Nonetheless, the question of how the controller should assess whether there is high risk to the data subjects in the event that they are not informed remains.

In order to try to answer this question two documents will be explored below; ENISA's recommendations on severity assessment of data breaches (ENISA, 2013) and WP29 Guidelines on data breach notification (WP29, 2018). Both documents focus on data incident risk assessment with the objective of identifying the need to notify the competent supervisory authority. Nonetheless, the methodology and criteria proposed therein can provide the necessary guidance in outlining the risk assessment process in the case of Article 14.

WP29 Guidelines on data breach notification present a list of factors that should be taken into consideration when assessing the risk of a breach to individuals. From this list the following elements could be taken into consideration to assess the risk – and its severity – to the individual's rights provided when the information obligation cannot be met:

- Special characteristics of individuals (e.g. children or other vulnerable categories of data subjects);
- Nature, sensitivity and volume of personal data (e.g. sensitive data, combination of sensitive data or high volume of data);
- Special characteristics of data controllers (e.g. medical organisation);
- Ease of identification of individuals, and
- Severity of consequences for individuals (e.g. identity theft, physical harm, psychological distress, humiliation or damage to reputation etc.).

From these indicative elements the majority will only be met when the operational purpose of the AI or the objective of the controller involve training using sensitive data. Furthermore, as we have also mentioned above, in the datasets used for training, without prejudice to the specific characteristics of the controller and the AI, even if the data subject is identifiable, the process to identify them may be too difficult and most of the time of no interest to the entity training the algorithm. Similarly, except in cases where the training involves individual profiling the severity of the consequences to individuals will be limited.

Furthermore, the WP29 Guidelines refer to ENISA's methodology in assessing severity of breaches. Below, we will investigate whether the methodology can also be deployed in the objective assessment of the controller under Article 14.

ENISA is following a three (3) criteria-based methodology to assess the severity of personal data breaches:

- Data Processing Context ("DPC") which evaluates the type of datasets at stake, together with a number of factors linked to the overall processing;
- Ease of Identification ("EI") which determines how easily can the identity of individuals be deduced; and
- Circumstances of breach ("CB") which addresses the specific circumstances of the breach, including mainly the loss of security of the breached data, as well as any involved malicious intent.

The final score of the severity assessment is extracted using the following formula:

$$\mathbf{SE = DPC\ x\ EI + CB}$$

Based on this final score, four levels of severity are defined:

| | | |
|---|---|---|
| $SE < 2$ | **Low** | Individuals either will not be affected or may encounter a few inconveniences, which they will overcome without any problem (time spent re-entering information, annoyances, irritations, etc.). |
| $2 \leq SE < 3$ | **Medium** | Individuals may encounter significant inconveniences, which they will be able to overcome despite a few difficulties (extra costs, denial of access to business services, fear, lack of understanding, stress, minor physical ailments, etc.). |

| | | |
|---|---|---|
| $3 \leq SE < 4$ | **High** | Individuals may encounter significant consequences, which they should be able to overcome albeit with serious difficulties (misappropriation of funds, blacklisting by banks, property damage, loss of employment, subpoena, worsening of health, etc.). |
| $4 \leq SE$ | **Very High** | Individuals may encounter significant, or even irreversible, consequences, which they may not overcome (financial distress such as substantial debt or inability to work, long-term psychological or physical ailments, death, etc.). |

**Table 3** – ENISA's Severity Score

From the ENISA criteria only CB cannot be used as such for our assessment (ENISA, 2013, p. 19-20). On the contrary, DPC and EI can be used, to also assess the risks for the data subjects.

Detaching the criteria from the data breach scenario, in reality DPC assesses the nature and volume of data used and EI the identifiability risk for the data subjects. In other words, DPC and EI are assessing part of the necessary elements that the controller needs to take into consideration as part of their objective assessment of the risks involved for the data subjects, as identified in Recitals 62[12], 75[13] and 76[14].

In order to meet the remaining criteria referred in the recitals (e.g. safeguards that the controller has adopted and purposes of processing) CB can be adapted to Circumstances of Processing ("CP") which would address the specific characteristics of the processing, including mainly whether the controller has put in place suitable safeguards, such as technical and organisational measures.

The ENISA levels of severity will remain as such with a change in the interpretation; when the severity score is low or medium the controller can benefit from the exemption of Article 14.5. lit. b. and proceed with processing after making the information publicly available on their website. When the severity of the risk is high or very high it will be safe to assume that the controller cannot benefit from the exemption. Each analysis should be well documented and

[12] *i.e.* number of data subjects involved and the age of the data.

[13] *e.g.* whether the data subjects might be deprived of their rights and freedoms or prevented from exercising control over their personal data.

[14] *i.e.* nature, scope, context of processing.

supported to allow the controller to present it in case they are requested by a competent supervisory authority or court.

In the light of our analysis above a possible risk assessment methodology and severity assessment formula for the conditions of the exemption of Article 14.5.lit.b. could be the following:

$$\mathbf{SE = DPC \times EI + CP}$$

### *DPC Calculation*

For the DPC calculation the ENISA classification of personal data and their scoring can be used (ENISA, 2013, p. 15-16). Depending on the category of data (simple data[15], behavioural data[16], financial data[17], sensitive data[18]) and specific aggravating factors such as the volume of data being processed, the characteristics of the controller (e.g. hospital developing AI-assistant) or of the individuals (e.g. children or other vulnerable individuals) included in the datasets, the DPC scores may vary between a preliminary basic score (also resulting when several decreasing severity factors exist) of 1 to 4 when various aggravating factors exist.

More specifically, ENISA scores simple data with a preliminary basic score of 1, behavioural data with 2, financial data with 3 and sensitive data with a basic score of 4. At the same time, the nature, the age of data as well as the source of data should also be taken into consideration either as aggravating or decreasing factors. For example, while data relating to political affiliations are sensitive data with a preliminary basic score of 4 if, for example, the data was made publicly available by the data subject the DPC score of this data can be decreased to 1. Similarly, the relevance and age of data can also impact its DPC score. For instance, if the income of an individual from 35 years ago is processed it does not pertain the same risks as more recent financial information.

### *EI Score*

The EI score will depend on the type and volume of data of the same individual being processed and whether this or the combination of this data facilitates the identification of the individual by the controller.

In this regard the score will vary between 0,25 when, for instance, the controller has no additional information for the data subject (e.g. part of the email

[15] *e.g.* biographical data.
[16] *e.g.* location, traffic data, preferences etc.
[17] *e.g.* income, transactions, credit cards etc.
[18] *e.g.* health, political views and affiliations etc.

address) to 1 when, for instance, the data available reveal the individual's identity (e.g. ID card with all the data). As an interim score of EI 0,75 is set when the data that is available to the controller can reveal significant information for the data subject (e.g. social security number with date of birth) or can easily be linked with other information (e.g. email address linked with full name and data of birth) (ENISA, 2013, p. 17-18).

As it can be understood, the EI score is quantifying the intrinsic higher risk for the individuals when they are easily identifiable, while at the same time it reduces the chances of the controller benefitting from the exemption of Article 14. In other words, the highest the EI score the less effort would need to be exhibited by the controller in order to provide the necessary information to the data subjects.

### *CP Score*

Lastly, in relation to CP a scoring system between 0 to 0,5 can be used where the characteristics and objectives of the processing (e.g. training for profiling purposes and combination of various datasets) and the characteristics of the controller (e.g. hospital developing AI-assistant) as well as the existence and nature of any technical and organisational measures deployed (e.g. state-of-the-art encryption) will be taken into consideration.

For the calculation of the CP, further guidance can be provided by the recently published report of the Spanish Data Protection Authority (Agencia Española Protección Datos "AEPD") (AEPD, 2021) on auditing processing activities involving AI. According to AEPD, in order to meet the objective of transparency and the obligation of providing information about the processing procedure to data subjects, both the data source and the logic of the AI should be "accessible, understandable and can be explained" (AEPD, 2021, p. 14). Amongst the proposed criteria in order to assess whether the objective has been achieved a series of measures are referred:

- "Documentation of data sources and implementation of an information mechanism;
- The characteristics of the data used to train the AI are identified, documented and duly justified;
- Selection and adoption of state-of-the-art methods that can facilitate readability, logic comprehension, internal consistency and explainability of the AI, including deep learning AI

- Information regarding AI metadata[19], its logic and consequences that may arise from its use is accessible to data subjects together with the appropriate means to exercise their rights;
- The logic of the AI is well documented while its behaviour regarding input datasets, data use, intermediate data and output data are traceable;
- (Mechanisms to mitigate the risks to the data subjects due to an erroneous behaviour of the AI have been established".[20]

Through this list of criteria, it is clear that the training process is in itself essential. Therefore, pursuant to our analysis, aside from suitable technical and organisational mechanisms, even if the data controller is not able to provide directly to the data subjects the necessary information and deliver the information notice, a thorough record of characteristics of the datasets and the data used for the training of AI should be carried, including their sources or the characteristics of their sources and the selection process of the sources (AEPD, 2021, p. 22).

Thus, the CP score will be 0 when the controller has established sufficient safeguards and has ensured that information available on their website are suitable, meaning including information on the processing, the logic, and objective of the AI which is easily accessible to individuals and has kept a thorough documentation of the datasets used and their sources. An interim score of 0,25 will be given when there is no documentation of the data sources and of the processing, however state-of-the-art technologies have been deployed that can ensure the protection of data and the readability of the AI logic. Lastly, a score of 0,5 will be given in situations where no, or not sufficient, measures have been taken and no information is made available to the data subjects. This aggravating element not only translates to higher risks for the rights of individuals but also to non-compliance with the Article 14.5 lit. b in fine, which requires the controller to put in place appropriate measures to protect the data subject's rights, including making the information publicly available.

The risk-assessment exercise described above can be carried independently or as part of a DPIA. As a proactive and objective analysis, it can support the controller's decision and processing activities in the operational stage of AI. Of course, in the event that this evaluation is presented to the competent data protection authorities, the authority will be free to evaluate its result and carry their own assessment.

Lastly, independently or together with a DPIA and the risk assessment exercise described above, an assessment regarding the trustworthiness of its algorithm in accordance with the Assessment List for Trustworthy Artificial

[19] Metadata of AI are parameters used in the training process.

[20] (AEPD, 2021, p. 14)

Intelligence of the High-Level Expert Group on Artificial Intelligence ("AI HLEG") (AIHLEG, 2020). Note that the Assessment List also addresses privacy and data governance aspects of AI (AIHLEG, 2020, p. 12-13). Considering that the European Commission's Proposal for the regulation on AI Act is under discussion, showcasing compliance with the proposed self-assessment can assist in futureproofing the algorithm.

## Additional Measures

Aside from the risk assessment that should be carried by the controller, the controller "*shall take appropriate measures to protect the data subject's rights and freedoms and legitimate interests, including making the information publicly available*" (Article 14.5. lit. b in fine GDPR). This means that the information on the processing shall be made available on the website of the controller and appropriate technical and organisational means, such as state of the art pseudonymisation techniques shall be deployed to support compliance with the data protection by design and by default ("DPbDD") measures as well as with the security obligations of Articles 25 and 32 GDPR, respectively. In other words, these technical and organisational solutions are necessary to comply with the transparency obligations of Article 14 GDPR.

More specifically, pursuant to Article 25 the controller is responsible to comply with the DPbDD obligations outlined therein. From a technical standpoint they aim to encourage controller to make use of state-of-the-art certifications and codes of conduct. It is essential that DPbDD measures are considered early in the lifecycle of the technology. In relation to AI, this means that the controller should implement DPbDD measures already when designing the algorithm as well as the processing activities, such as the use of datasets for the training of AI. For older already existing algorithms, that were designed before the application of the GDPR, the DPbDD obligations still apply and the existing systems require review to ensure compliance with Article 25.

GDPR does not identify specific technical measures. Nonetheless, there is a clear preference for encryption, since it is explicitly mentioned in three Articles of GDPR, Namely, Articles 6, 32 and 34 GDPR, as an example of a suitable technical measure to ensure data protection and security.

Moreover, in accordance with the EDPB Guidelines on DPbDD the technical and organizational measures and necessary safeguards can be understood in a broad sense, meaning any method or means that a controller may deploy in the processing (EDPB, 2020, p. 6). The means used need to be appropriate, meaning that the measures can ensure the effective application of the data protection principles, and in the light of Article 14, the principle of transparency. The requirement to appropriateness is thus, closely related to the requirement of effectiveness.

The EDPB Guidelines present various indicative measures that may be relevant for controllers designing and training algorithms such as pseudonymisation; storing personal data available in a structured, commonly machine readable format; providing information about the storage of data; training employees about basic "cyber hygiene"; establishing privacy and information security management systems, obligating processors contractually to implement specific data minimisation practices, undertake certification, follow best practices and codes of conduct and erase or anonymise the data when they are on longer useful for the training or validation purposes (e.g. outdated, not valid etc.) etc. (EDPB, 2020, p. 6).

In particular regarding pseudonymisation[21], Recital 28 states that the application of pseudonymisation techniques "to personal data can reduce the risks to the data subjects and help controllers and processors meet their data-protection obligations". The most evident benefit of pseudonymisation is to "mask" the identity of the data subjects from third parties. Additionally, pseudonymisation can contribute towards the principle of data minimisation when the controller does not need to have access to all the information of the data subjects as well as towards data accuracy since it can support data preservation. Moreover, taking into account the need for high levels of data utility in the context of AI-training, verification and validation, it is expected that in the future, and in particular in relation with AI-related and Big Data analytics processing activities, the use of pseudonymisation as one of the main technical solutions for data protection will only increase (ENISA, 2021, p. 49).

Since, as underlined above, GDPR does not identify specific techniques that meet its requirements, ENISA on its report on pseudonymisation techniques and best practices provides further guidance on the selection and deployment of an efficient pseudonymisation technique (ENISA, 2019). In this regard, ENISA suggests that to ensure a robust pseudonymisation process, capable of reducing the risks of identification of the data subjects while preserving the necessary utility for the use of the data, a high level of technical competence is required by the data controllers. The difficulty of selecting a suitable pseudonymisation technique only rises in the context of AI training datasets. Therefore, ENISA proposes a risk-based approach test to assess whether a technique is robust and

[21] Under Article 4(5) GDPR, pseudonymisation is defined as *"...the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person".*

effective in the scope of GDPR. Such assessment should take into consideration the level of protection granted by the particular technique, the overall context of the specific processing, the utility and scalability levels that the controller is aiming for and, lastly, the state-of-the-art (ENISA, 2019, p. 42-43). It is important to note that in the same report, ENISA suggests that in order to achieve higher levels of data utility and protection, perhaps, a combination of different techniques is advised, especially taking into account that there are currently manifold techniques with different characteristics.

Furthermore, EDPB identifies a series of examples of specific technical and organisational measures appropriate for each processing principle. In relation to the principle of transparency, EDPB highlights that any selected measures should support not only the principle itself but also the implementation of the Articles where it's embedded, including Article 14. Although a listing of the key design and default elements for the principle of transparency as identified by the EDPB exceeds the objective of our analysis, it is interesting to refer specifically to the importance of multi-channel information. Multi-channel information proposes the use of different channels and media to "increase the probability for the information to effectively reach the data subject". The importance and utility of this DPbDD is, thus, evident when the controller does not have direct contact with the individuals.

Lastly, as stated in the EDPB Guidelines on DPbDD, privacy-enhancing technologies ("PETs") (e.g. access control, SSL/TLS encryption, VPNs etc.) that have reached state-of-the-art maturity can be employed as a measure in accordance with the DPbDD requirements if appropriate in a risk-based approach. This is important for the assessment and ultimately selection of efficient and robust measures by the controller that can be used both for the designing of the algorithm as well as during the AI training process. Notwithstanding, PETs in themselves do not necessarily meet the DPbDD requirements. The appropriateness and effectiveness of PETs should be subject to an objective assessment by the data controller. Such assessment, pursuant to the principle of accountability, should be thoroughly documented. For instance, this documentation can be part of the risk-assessment proposed above or of the DPIA.

Moreover, suitable organisational measures to mitigate the risks for individuals when is not possible to send information notice should also be deployed. Similarly, to the technical measures, the organisational measures that can be deployed vary. For instance, monitoring internal activities and processes, or requesting contractually, when possible, from the third-party providers of the datasets who collect directly the data from the data subjects to provide sufficient information to the individuals which can cover also the processing activities to be undertaken by the controller for training the AI (ZANFIR-FORTUNA, 2020, p.

444). Of course, in this case the exemption of Article 14.5 lit. a. for situations where the data subject has already been informed and identified will apply.

Although the contractual enforcement of such an obligation to the third-party dataset providers will allow the controller to demonstrate, if requested by a supervisory authority or court, that the necessary information was provided to the data subjects at the time of data collection and they are not required to provide additional information, it may prove complicated as it presupposes, firstly, a formal contractual relationships with the third party controller that will permit contract negotiation and, secondly, a certain level of market power of the AI-controller over the third-party provider. Therefore, the scope of application of such measure may be limited in the AI field and other alternatives can be explored.

In addition, the importance of adequate technical and organisational measures was recently highlighted in the context of international data transfers, following ECJ's Schrems II ruling (C-311/18)[22]. Shortly after, EDPB adopted Recommendations 01/2020 on measures that supplement transfer tools to ensure compliance with the EU level of protection of personal data when data is transferred outside of the European Economic Area (EEA) and where there is not an adequacy decision in place (EDPB, 2020). Considering that most of the times that controller wishing to use certain datasets for the training, verification and validation process of AI is not in a position to know where is the data stored as well as that many data sharing providers are based outside of the EEA, it is safe to assume that there may be international data transfers involved. Therefore, taking into account the limited current number of adequacy decisions[23], it is necessary for organisations to identify and adopt supplementary measures to ensure an adequate level of protection when the data is in transit and at rest and throughout the processing operations. According to the EDPB Recommendations, the supplementary measures, although they should be assessed on a case-by-case basis for each international transfer, may be contractual, technical or organisational or a combination of the above (EDPB, 2020, p. 15).

Lastly, it should be noted that the analysis above has also an impact on controllers and processors established outside of EEA. Without prejudice to public international law provisions that may stipulate the application of GDPR, GDPR also applies to processing of data belonging to individuals who are in the EU even if the controller or processor is established outside of the EU, when these

[22] In its C-311/18 Schrems II case, ECJ invalidated the EU-US Privacy Shield agreement, based on which, to a large extent, the transfers of data to the US were performed.

[23] The European Commission has the power to determine, on the basis of Article 45 GDPR whether a country outside the EU offers an adequate level of data protection. At the time of running there is an adequacy decision for the following countries: Andorra, Argentina, Canada (commercial organisations), Faroe Islands, Guernsey, Israel, Isle of Man, Japan, Jersey, New Zealand, Switzerland and Uruguay.

processing activities are related to the offering of goods or services or the monitoring of the individual's behaviour. Considering that the organisations processing the datasets are not always in a position to determine whether personal data of EU-based data subjects is included in the datasets as well as the broad language of Article 3 regarding GDPR's territorial scope, it is safe to assume that the likelihood of GDPR applying is high.

In the light of the above, the globalised and highly connected world of Big Data, together with the extraterritorial reach of GDPR, present manifold opportunities for SMEs established in large developing countries that have access to AI-technology and talent. Taking into account the strategy of the European Commission to strengthen and promote European AI initiatives, including through synergies, a proactive adoption of technical and organisational measures in compliance with GDPR can promote future collaborations with EU-based entities, intensify international research efforts and facilitate data transfers and exchange of information in the field. Similarly, from a B2C point of view, a voluntary compliance with GDPR and adoption of efficient mechanisms, in particular, pursuant to the transparency principle, can significantly increase consumer trust to the AI-based products and services offered.

## CONCLUSION

As every exemption to or restriction of the rights of the data subject, the exemption of disproportionate effort of Article 14.5.lit.b is narrowly interpreted and requires a high degree of responsibility on behalf of the controller.

In our analysis, we aimed to design a clear risk-assessment methodology that can be carried by controllers who think that are not able to communicate the necessary information to the data subjects but wish to ensure their data protection compliance.

The lack of guidance on the matter by EDPB or other supervisory authorities, creates a level of uncertainty regarding the steps that a controller wanting to use large datasets to train their machine or deep learning algorithm should take. This uncertainty may hamper the uptake or AI and the activities, especially, of SMEs that may not have the means or market position to select or monitor the third-party dataset providers nor sufficient data protection expertise.

By adapting the available recommendations of ENISA and EDPB on risk assessment in the context of DPIAs or of data incidents to the needs and requirements of Article 14, a higher level of flexibility is achieved while using already successful formulas that are known to controllers.

Without prejudice to any additional obligations that may arise due to the nature of the processing or due to the characteristics of the controller and of the data subjects, aside from an objective and clearly documented balancing test, certain additional measures should be taken to mitigate the risks for the data subjects. Firstly, the controller is required to make publicly available in a clear and legible manner the information regarding the processing, as well as the information regarding the automated-decision making. Secondly, the controller should also deploy suitable technical and organisational measures that demonstrate their compliance not only with the obligations of Article 25 and 32 but also with Article 14.5 lit. b. and the transparency principle. In this way, the security of the data throughout the AI lifecycle can be promoted. Lastly, it is recommended that a self-assessment based on the Assessment List for Trustworthy AI of the AI HLEG.

## REFERENCES

AEPD. *Audit Requirements for Persohnal Data Processing Activities Involving AI.* Agencia Española Protección Datos, 2021.

AIHLEG. *The Assessment List for Trustworthy Artificial Intelligence (ALTAI) for self-assessment.* High-Level Expert Group on Artificial Intelligence, 2020.

BALDWIN, R.; CAVE, M.; LODGE, M. (Eds.). *The Oxford Handbook of Regulation.* Oxford: Oxford University Press, 2010.

BÖCKENFÖRDE, E.-W. *Escritos sobre derechos fundamentales.* Tradução de Juan Luis Requejo Pagés e Ignacio Villaverde Menéndez. Baden-Baden: Nomos, 1993.

CABRAL, T. S. AI and the Right to Explanation: Three Legal Bases under the GDPR. In: HALLINAN; LEENES; DE HERT *Data Protection and Privacy:* Data Protection and Artificial Intelligence. Oxford, UK: Hart Publishing, 2021.

CARLSSON, U. The Rise and Fall of NWICO: From a Vision of International Regulation to a Reality of Multilevel Governance. *Nordicom Review*, v. 2, p. 31-68, 2003.

EDPB. *Guidelines 4/2019 on Article 25 Data Protection by Design and by Defualt.* European Data Protection Board, 2020.

EDPB. *Recommendations 01/2020 on measures that supplement transfer tools to ensure compliance with the EU level of protection of personal data.* European Data Protectin Board, 2020.

ENISA. *Recommendations for a methodology of the assessment of severity of personal data breaches, Working Document.* European Union Agency for Network and Information Security, 2013.

ENISA. *Pseudonymisation techniques and best practices: Recommendation on shaping technology according to data protection and privacy provisions.*

European Union Agency for Network and Information Security. Athens, Greece, 2019.

ENISA. *Data Pseudonymisation: Advanced Techniques and Use Cases*. European Union Agency for Network and Information Security. [s.l.]. Athens, Greece, 2021.

ERK, J. Austria: A Federation without Federalism. *Publius*, v. 34, n. 1, p. 1-20, 2004.

FRA. *Handbook on European data protection law.* European Union Agency for Fundamental Rights, 2018.

GEORGIEVA, L. Article 11 Processing which does not require identification. In: KUNER, et al. *The EU General Data Protection Regulation (GDPR) – A commentary*. Oxford, UK: Oxford University Press, p. 391-397, 2020.

HÄBERLE, P. *Die Wesensgehaltgarantie des Art. 19 Abs. 2 Grundgesetz.* Karlsruhe: C.F.Müller, 1962.

HUMBOLDT, W. V. *On Language:* On the Diversity of Human Language Construction and its Influence on the Mental Development of the Human Species. Tradução de Peter Heath. Cambridge: Cambridge University Press, 1999.

ICO. *Guide to the General Data Protection Regulation (GDPR)*. Information Commissioner's Office, 2018.

KINDYLIDI, I. Smart Companies: Company & board members liability in the age of AI. *UNIO – EU Law Journal*, v. v. 6, n. n. 1, p. 115-141, 2020.

KROLL et al. Accountable Algorithms. *University of Pennsylvania Law Review*, v. 165, 2017.

LEVY, B.; SPILLER, P. *Regulations, Institutions and Commitment.* Cambridge: Cambridge University Press, 1996.

LUHMANN, N. *Law as a Social System.* Tradução de Klaus A. Ziegert. Oxford: Oxford University Press, 2004.

MALGIERI; COMANDÉ. Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation. *International Data Privacy Law*, v. 7, n. 3, 2017.

PRICE, M. E.; NOLL, R. G. *A Communications Cornucopia:* Markle Foundation Essays on Information Policy. Washington, DC: Brookings Institution Press, 1998.

ROSE-ACKERMAN, S.; LINDSETH, P. L. (Eds.). *Comparative Administrative Law.* Cheltenham, UK: Edward Elgar, 2010.

SARTOR, G.; LAGIOLA, F. *The impact of the General Data Protection Regulation (GDPR) on artificial intelligence.* European Parliamentary Research Service - Scientific Foresight Unit (STOA) - Panel for the Future of Science and Technology, 2020.

WACHTER; MITTELSTADT; FLORIDI. Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation. *International Data Privacy Law*, v. 7, n. 3, 2017.

WP29. *Opinion 03/2013 on purpose limitation.* Article 29 Data Protection Working Party, 2013.

WP29. *Guidelines on Data Protection Officers ('DPOs').* Article 29 Data Protection Working Party, 2016.

WP29. *Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679.* Article 29 Data Protection Working Party, p. 22. 2017.

WP29. *Guidelines on transparency under Regulation 2016/679.* Article 29 Data Protection Working Party, 2017.

WP29. *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679.* Article 29 Data Protection Working Party, 2018.

WP29. *Guidelines on Personal data breach notification under Regulation 2016/679.* Article 29 Data Protection Working Party, 2018.

ZANFIR-FORTUNA, G. Article 13. Information to be provided where perosnal data are collected from the data subject. In: KUNER, et al. *The EU General Data Protection Regulation (GDPR) – A Commentary.* Oxford, UK: Oxford University Press, p. 413-433, 2020.

ZANFIR-FORTUNA, G. Article 14. Information to be provided where personal data have not been obtained from the data subject. In: KUNER, et al. *The EU General Data Protection Regulation (GDPR) – A Commentary.* Oxford, UK: Oxford University Press, p. 434-448, 2020.

**The Law, State and Telecommunications Review / Revista de Direito, Estado e Telecomunicações**

***Contact:***
Universidade de Brasília - Faculdade de Direito - Núcleo de Direito Setorial e Regulatório
Campus Universitário de Brasília
Brasília, DF, CEP 70919-970
Caixa Postal 04413

***Phone:*** +55(61)3107-2683/2688

***E-mail:*** getel@unb.br

Submissions are welcome at: https://periodicos.unb.br/index.php/RDET