aepd agencia española protección datos

Technologies and Data Protection in Public Administrations

November 2020

EXECUTIVE SUMMARY

Technology is changing the way Public Administrations work together with their relationship with citizens in search of better timings, better accessibility, simpler procedures and less cost. On the side, a specific risk exists associated to the processing of personal data by Public Administrations when they implement product and services using emerging technologies.

In this regard, this document seeks to perform an analysis of some of these technologies applied by Public Administrations, so as to highlight their most relevant features from a data protection perspective and to bring to light some of the inherent risks associated to their use in any given processing activity.

This document does not aim to be a guideline or a set of regulating criteria. With that in mind, the Spanish Agency of Data Protection has been publishing specific guidelines on some of the aspects that will be addressed below, as well as some insights on the GDPR or the specific domestic legislation implementing the GDPR. Reference will be made to such documents throughout this text. The purpose of this document is to highlight some of the aspects that could most define the use of these technologies in processing activities by Public Administrations.

Therefore the list described of fulfilment guarantees and risks to be managed is not an exhaustive one, but rather a list of the basic aspects that need to be taken into account in any processing activity that involves such technologies. The list of technologies studied is just a first approach. The content of this document is open and seeks to be completed through further issues that will contain the feedback of the readers and we plan to extend it to other specific technologies for the purposes of setting the risk control objectives associated to each of them.

This document is mainly addressed to Data Protection Officers of the Public Administrations and civil servants in charge of promoting, managing and using these technologies within the Administration, albeit the content, and, most of all, the references included in it may likewise be useful to a broader audience, such as managers in private companies that may acts as data processors or developers for the Public Administrations as well as the citizens themselves, so as to understand the way in which these technologies affect them within the framework of the services provided to them by the Public Administrations.

A simple wording has been sought in order to draft this document, far from the formalities both at technical and at legal level, so as to serve by way of introduction to those who are in charge of designing procedure activities that include emerging technologies within the public sector.

Keywords: GDPR, public administrations, eAdministration, Digital Administration, risks, data protection, privacy, Big Data, blockchain, cookies, artificial intelligence, social networks, smart cities, technology.

ASTIC (Asociación Profesional de Cuerpos Superiores de Sistemas y Tecnologías de la Información de las Administraciones Públicas-Profesional Association of Higher Bodies of Information Technologies and Systems of the Public Administrations),the Data Protection Officer of the Parliament of Andalucía, the Election Board of Andalucía and the Ombudsman for the Andalusian People, Mr Iñaki González-Pol and Mrs. Sara Degli Esposti, a researcher at the Instituto de Políticas y Bienes Públicos of the CSIC (Institute of Public Properties and Policies of the Higher Centre of Scientific Investigation) have participated in the review of this document.

CONTENTS

I. INTRODUCTION

A. TECHNOLOGY AND PUBLIC SERVICE

Technology is part of our lives. It surrounds us and makes our life easier while preserving and even enriching our rights and freedoms. A great part of our personal and professional life involves interacting through computers and mobile phones with a technology that makes us need to renew ourselves and create new skills on an ongoing basis.

Technology has made possible thing that were unimaginable in the past. We can make purchases from home, speak as we see people that are thousands of miles away from us, manage much more complex tasks and manage them much faster than before. However, it is this very technology that serves as an ally to make our daily life easier that can also be a threat to our rights and liberties if not correctly understood. Our dependence on machines and their good working is virtually complete. Never in humankind history had so many people who knew about us, about our tastes, our traditions or secrets, and never had a person sitting before a computer held so much power in their hands to help other or to damage them.

The use of information and communication technologies or ICTs by the public sector, the Digital Administration, allows to offer more advanced mechanisms in order to implement both the services provided to us and the services needed for their internal functioning. Such services often are personal data processing activities that are implemented through the use of one or several technologies thus providing efficacy, efficiency, availability, interoperability, and rationalisation of resources, among many benefits.

Public managers, in charge of personal data processing activities, are under the obligation to fulfil the existing regulations on personal data protection within their jurisdiction, and the technical and organisational measures needed, as the case may be, to guarantee, inter alia, the rights of citizens, the transparency, and the application of the principles of proactive responsibility, as well as the other principles with regard to processing activities contained in the General Data Protection Regulation (GDPR) must be applied.

Consistency of a processing activity with the GDPR requires for every step in the design, the implementation and the management of the processing activity to include all guarantees from the beginning. Simply put, the following questions need to be answered: “who?” (Articles 24 and 26 on responsibility), “for what purpose?” (Article 5.1.b GDPR. Principle of purpose limitation), “why?” (Article 5.1.a Principle of legitimacy and loyalty and the rest of Chapter II), “how much?” (Article 5.1.c Principle of minimisation), “when?” (Article 5.1.e Principle of limitation of the preservation period). Then will come the “how?” question (Article 5.2 Principle of proactive responsibility, which includes, at the same time, 6.1.d and e, with chapters IV and V) including the question “where?” (Articles 28 on the data processor, 29 y chapter V) and the question “in what way?” regarding the rights of data subjects that need to be preserved (Article 5.1.a Principle of transparency and Chapter III).

The possibility could exist of having drafted a formal and exhaustive guideline from a technical point of view and from a legal point of view regarding data protection to address all these questions. If such had been the case, a more complete and useful document would have been obtained for people specialised in technology and data protection. However, the decision has been adopted to create a document aimed at readers who do not know, nor wish to know, the insights of these technologies, but rather, they wish to learn how they work and how they can affect data processing activities. The simple wording in this document has been chosen with them in mind, for the sake of clarity over detail, so that readers are offered a horizontal vision of technologies and their risks. However, links and references have been added in the foot of the pages for those in need of learning more about some of the ideas addressed in this document so that they can complete their information through articles, recitals and complementary documents of interest.

B. THE RISK FOR RIGHTS AND FREEDOMS IN PUBLIC ADMINISTRATIONS

The General Data Protection Regulation and the Organic Act on Personal Data and Guarantee of Digital Rights (hereinafter, the LOPDGDD) envisage the possibility that the introduction of disruptive, innovative or not sufficiently evolved technologies in personal data processing activities be a factor that may increase the risk towards the rights and freedoms of data subjects, a risk that needs to be assessed.

The assessment, management, and minimisation of the risk for the rights and freedoms is an obligation by the data controller (Articles 23.2.g, 24.1, 25, 32, 33, 34, 35 and 36, inter alia) and is part of the list of regulation compliance. Even if the GDPR contains certain guidelines, it is not precise at the time to identify and regulate how to management the risk of each processing activity specifically.

Any processing implemented within the frame of the Public Administrations entails a series of risks that need to be managed as any other process within the organisation. This management does not differ from the management that needs to be performed at any other entity within the frame of activities that imply the processing of personal data. By way of an example, at the time to launch a service or a product, all entities need to manage, inter alia, the financial risk of addressing and implementing a new initiative, the risk of the cost that entails the relation with the benefit it provides, the risk that the new service may be deployed according to a schedule, the risk of the opportunity cost, the risk of the reliability of the technical options or technologies to be used, the legal risks of any future modifications to the legislation, the risk of fulfilment that exposes such entity to administrative sanctions, civil sanctions or criminal sanctions, the environmental risk, the security risk with regard to the continuity of the system, the risk of reputation attacks, the risk of fraud, etc. All these risks are not analysed individually, but rather, they must be analysed as a whole [1] in order to reach a decision within the frame of a holistic approach that takes the context into consideration.

As part of this comprehensive analysis we may find the management of the risks for the rights and freedoms of individuals. Such an assessment is not related to the risk of fulfilment of the principles of processing, the rights and the obligations established in the GDPR, that is to say, through the assessment of the possibility of incurring in sanctions, significant financial losses or loss of reputation as a result of a breach of the laws, the regulations, the internal rules and the codes of conduct.

The management of the risk for the rights and freedoms within the frame of personal data processing activities needs to envisage the possibility that, even with a formal compliance of the provisions in personal data protection regulations, the context and the scope in which such processing activity is taking place may introduce a certain degree of uncertainty in terms of its necessity and its proportionality, as well as of the efficacy and the effectiveness of the legal and technical guarantees applied to it.

It must be taken into account that the processing of personal data by the Public Administrations implies risks that are different from the risks arising out of a processing carried out by any other data controller. such risks rise, at least, as a result of the volume of data subjects affected the extent of the data collected, the impossibility to oppose to the processing, in any cases, and the power or inherent asymmetry existing between the Public Administrations, the citizens or the data subjects whose data are being processed. On another note, regardless of the fact that all processing activities carried out by the Public Administrations are guided by a spirit of public service, a spirit that also inspires the work of

[1] ISO standard 31000:2018: Risk Management-Guidelines

their employees, these possible risks could materialise over the citizens in certain situations such as, for example, situations of breaches of the rule of law, situations of abuse by the public employees, circumstances of massive leaks or selective leaks of personal data as a consequence of security breaches, in the event of a legislation modification including modifications in third countries to which the data have been transferred, in the event of corruption, emergency situations that are beyond control, etc.

As a consequence, the Public Administrations, to the extent that they are data controllers of the data of the citizens, before implementing new processing activities or modifying the services already provided that use new technologies, must identify the risks to which such processing would be exposed. The technical and organisational measures should likewise be adopted that are needed so that, by design and by default, such measures allow to suppress or to, at least, mitigate the damages, up to an acceptable level, that may arise out of the processing for the rights and freedoms of citizens.

It must be taken into account that the citizens that could encounter this situation could not only be the citizens that are subject to administration but also the civil servants themselves while performing their duties.

It is obvious that there is no zero risk. Both ignoring this risk and acknowledging it to reduce its impact or the probability it may occur will not make it disappear. Risk is not static, and it evolves continuously as a consequence of the evolution of its background. This is why, once this risk is identified, the data controller must continuously supervise and manage such a risk. A correct attitude is that of acknowledging the risks, assessing its consequences, taking measures in order to minimise it, and control its effectiveness in a changing context. This ongoing supervision scheme is what is defined as risk management.

C. STRUCTURE OF THE DOCUMENT

The following sections review a series of technologies that have been added by the Public Administrations as means and formats for several personal data processing activities. More precisely, the document focuses on the analysis of some specific aspects regarding compliance and the risks that may appear in processing activities [2]due to the use of these technologies.

- Cookies and tracking technologies
- Social networks
- Cloud Computing
- Big Data
- Artificial Intelligence
- Blockchain and Distributed Ledger Technologies
- Smart Cities

In this edition, the selection has been limited to those technologies that are vastly used or have the potential to be implemented nowadays. New technological solutions may be added in further issues of this document that will most surely bring new risks with them that will need to be assessed.

A brief description, the goal of such a technology and its functioning is first provided for each of these technologies. Below, we have stated some of the specificities in terms of fulfilment whenever technologies are added to a processing activity within the frame of Public

[2] Apart from the risks detailed in this document, the controller will need to analyse the casuistry and the risks that are not specifically related to the technology that is being used with regard to the processing due to other circumstances as a result of the nature, the scope, the context and the purposes.

Administrations. It is important to highlight that such specificities of fulfilment are precise aspects that are highlighted with regard to the technology that is being used, and not with regard to the processing where they are being used, whose fulfilment requirements are considerably vaster.

Finally, some of the possible risks are listed for each technology that, in the same sense as the fulfilment requirements, are specific of each implementation of such technology, without making an assessment on the general risks associated to the specific processing activity on which they are being implemented. This last part would entail a global, complete, and comprehensive analysis that focuses on the processing and, therefore, more detailed and broader. On another note, it must be highlighted that some of the risks identified and the guarantees to approach them may be shared with other technologies and, consequently, the study performed and the conclusions reached could be extrapolated.

The aim of this document is to expose several aspects that are characteristic of these technologies regarding the protection of data when such technologies are used by the Public Administrations and that they seek to complement what has already been established by the GDPR, the national regulations and industry regulations together with specific guidelines already published such as the Guía y listado de Cumplimiento Normativo (Guideline and List of Regulation Compliance), Guía de protección de datos y Administración Local (Guide of data protection and Local Administration, Código de buenas prácticas en proyectos Big Data (Code of Conduct in Big Data projects), Guía para clientes que contraten servicios de Cloud Computing (Guide for clients hiring Cloud Computing services), Guía sobre el uso de Cookies (Guide on the use of cookies), **RGPD compliance of processings that embed Artificial Intelligence. An introduction**, **A Guide to Privacy by Design**, **Guidelines for Data Protection by Default,** Guía práctica para el Análisis de Riesgos (Practical Guide on Risk Analysis) and Guía práctica para la realización de Evaluaciones de Impacto en protección de datos (Practical Guide on the performance of Impact Assessments on data protection), **Guide on personal data breach management and notification**, etc. Their aim is to serve as a help tool and as a starting point for the analysis of compliance and management of the risk linked to the incorporation of technologies to processing activities that are performed within the scope of Public Administrations and are fully or partially supported by the technological solutions described in this document.

II. COOKIES AND TRACKING TECHNOLOGIES

A. WHAT ARE COOKIES?

Once they became broadly used on the Internet, the opportunity was discovered to know, for example, who was visiting your webpage or if it was the first time they were visiting it, from where they had landed on your website or which other pages in the portal had also been visited by the users, if they had registered or the possibility to offer them to save settings or log in details. *Cookies* allow you to implement these functions and more. For example, *cookies* allow advertisers to know what Internet sites are visited by users in order to offer them products according to their tastes. Those responsible for the portals may like to know as well the pages that have been visited so as to obtain statistics, adopt decisions, and a long et cetera.

The tracking of the activity by the users in the network is implemented through devices that are [3]generically called *cookies.* This is a strategy used by the servers to store and retrieve information of the user's device, which allows to process personal data when the user browses and interacts with the different applications and contents deployed in the Net. Indeed, under the denomination of *cookies* there are several techniques that allow to track the users both actively and passively, often in a way that is not very transparent. Among these tracking technologies we may find those that use the features of the device, the unique identifiers, and the user's browsing habits. Every time we ask for a page, an image or a content from a web server, we are conveying to such server at least, our IP address, which allows the server to know our geographical location[4], but also the browser model we are using and, therefore, our operating system, the device through which we are getting online and the state of update of such device. A web server may know if the user that is browsing has a pop-up blocker, how much storage is left in their device, what graphic card it uses, or how they move their mouse over the screen. All this information is grouped under the generic name of *fingerprint*[5] or digital fingerprint of the device, and may be used in Internet servers to link all the user's activity and thus create a profile of the user.

The GDPR makes a statement on *cookies* in Recital 30[6]acknowledging their capacity to create profiles on persons and identify them[7]. The special regulations that provide for the use of *cookies* for the services of the information society is contained in Section 22.2 of Act 34/2002 of 11 July on services of the information society and electronic commerce (LSSI in Spanish)*.* The AEPD criteria for the implementation of the regulations regarding *cookies* by service providers of the information society are contained in the Guideline on the use of cookies[8] that was published in 2019 and updated in July 2020. Such guideline includes as

[3] The term "dispositivo" appears in Spanish legislation as a translation of the English word "device" used by Directive 2002/58/EC or the Directive on ePrivacy. The semantic extent of the word "device" includes: artefact, strategy, gadget or resource.

[4] What is GeoIP and its benefits? https://serverguy.com/news/what-is-geoip/

[5] The Spanish Data Protection Agency- AEPD published back in September 2019 a guideline called Fingerprinting or Digital fingerprint of the device, accessible through https://www.aepd.es/sites/default/files/2019-09/estudio-fingerprinting-huella-digital.pdf

[6] The European Commission is working on a proposal of the Regulation of the European Parliament and of the Council concerning the respect for private life and the protection of personal data in electronic communications, better known as the ePrivacy Regulation Proposal https://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52017PC0010&from=ES, which would repeal 2002/58/EC

[7] Act 34/2002,on services of the information society and electronic commerce https://www.boe.es/buscar/act.php?id=BOE-A-2002-13758

[8] Available at https://www.aepd.es/media/guias/guia-cookies.pdf. Other authorities from neighbouring countries have likewise published guidelines recently, such as the British ICO(https://ico.org.uk/for-organisations/guide-to-pecr/guidance-on-the-use-of-cookies-and-similar-technologies/), the French CNIL (https://www.cnil.fr/en/cookies-and-other-tracking-devices-cnil-publishes-new-guidelines) or the Conference on German Data Protection Authorities https://www.datenschutzkonferenz-online.de/media/oh/20190405_oh_tmg.pdf. An interesting comparison may be found at https://iapp.org/resources/article/ico-and-cnil-revised-cookie-guidelines-convergence-and-divergence/. The guide on the use of cookies published by the AEPD has been updated in order to adapt it to the new Guidelines on consent reviewed by the European Data Protection Board in May 2020, with the mandatory implementation of these new criteria by 31 October 2020

well the definitions of the different types of *cookies*, the obligations of editors with regard to the information to be provided to their users, the way in which consent is collected prior to the use of cookies, and the responsibility of the parties at the time to use *cookies*. This document, with regard to the others, refers to such a guideline and the section below will address several specificities in the use of *cookies* by the Public Administrations.

B. COOKIES AND THE PUBLIC ADMINISTRATIONS

The applicability of the LSSI

Subjects obliged by the special regulation providing for *cookies* are service providers of the information society, understood as *"any natural or legal person that provides services, usually with a consideration, from distance, electronically and at the individual request of the recipient, as well as those services that are not remunerated by their recipients, to the extent that these constitute an economic activity for the service provider".*[9]*.*

As explained in the page on the Act on Services of the Information Society[10] of the Spanish Ministry of Economic Affairs and Digital Agenda:

> *In general terms, the LSSI is not applied to Public Administrations, as these do not have the quality of service providers of the information society defined in the annex thereof. Thus, certain activities that are typical of the Administrations, such as the electronic management of taxes or the information on services by a third party (such as the mere information in the webpage of a City Council on available country houses within a municipality) are considered as public activities or activities of public interest other than the "economic activity" referred to in the LSSI.*
>
> *Notwithstanding, when the activity of the Administration does have an economic character (such as, for example, the sale of tourism books by a dependent public entity of a City Council), the LSSI will apply.*

That is to say, under this approach, and as a general basis, a Public Administration with a portal web would not necessarily be considered as a service provider of the information society just for the mere fact of having a portal web, and, consequently, it would not be subject to the LSSI. It will solely be considered as such if it offers "*a service normally provided with a consideration, remotely, electronically, and at the individual request of the recipient*" for example, if a service is provided within a domain that pertains to a Public Administration and such service involves an economic activity, such as the sale of books, tickets, etc. In this last case, the LSSI will indeed apply to such administration for the precise domains where such economic activity is taking place.

On another note, the very definition of service of the information society states: "*the concept of the service of the information society also involves services that are not remunerated by their recipients, to the extent that these entail an economic activity for the service provider...*". A webpage typically contains third-party components, tools and applications such as social networks (Twitter, Facebook, etc.), content distribution channels (YouTube, Vimeo, etc.),or other services such as maps or translators that involve the use of tracking *cookies*, analysis or advertising. The fact of allowing the use of these *cookies* in exchange of the possibility of including such services within the web portal, could entail an economic activity provided such tools or applications are executed within the very domain of the Public Administration and are not links to third-party websites. Under these circumstances, the Public Administration Bodies who use third-party services implying the

[9] Act 34/ 2002, of 11 July, on services of the information society and electronic commerce, or LSSI.

[10] http://www.lssi.gob.es/la-ley/Paginas/preguntas-frecuentes.aspx?Faq=%C3%81mbito+de+aplicaci%C3%B3n

use of *cookies* will need to apply the obligations and guarantees established under the specific legislation and the AEPD guidelines.

Every time third-party cookies are included, even when means are provided that allow for the processing of personal data by third parties [11], either in the cases referred to above or other cases, such as allowing the use of third-party analytic cookies, the Public Administration must be diligent at the time to detect that these processing activities are taking place and at the time to guarantee that such processing activities comply with the information duty and the duty to obtain consent from the users.

In any event, when the Public Administration is not subject to the LSSI, provided that the processing activity carried out through the *cookies* involves personal data, such processing will be subject to the GDPR and all applicable personal data legislation. Notwithstanding, the Public Administrations, for the sake of transparency, cold share information on the *cookies* that are being used when the duty to obtain consent is not applicable. In order to address the relevant obligations, the Public Administrations may use the directions described in the Guideline on the use of cookies published by the AEPD and referred to above.

Use of cookies in official portals and apps

The web portals of the organisations are not, in general terms, a set of static pages with a fixed content, but rather, they generally are supported by a CMS or Content Management System[12]). These managers are executed in the web server and build, in a dynamic way, the pages that users request from it out of content pieces stored in a data base. There are portals that use customised managers, but most of them use commercial programs or open-code programs. These managers can be using *cookies* for their functioning in such a way that the official portal may request the use of *cookies* as a technological prerequisite. The use of *cookies* may extend as more *plugins* are added to increase the functionality and the features of the CMS. The same happens with the apps or mobile applications that are distributed by some Public Administrations. Usually, such apps are built on a commercial programming environment or open-code environment[13], and they normally include third-party SDKs[14] that allow to increase functionality in a smooth and fast way, but that may be including the use of additional *cookies*[15] at the same time or other tracking techniques, such as the processing of unique advertising identifiers[16].

In view of these circumstances, the controller of the official portal must obtain information and, as the case may be, verify what type of *cookies* are requested by the content manager together with the *plugins* that are to be used. As the case may be, the controller needs to adopt the necessary measures to suppress those that are unnecessary for the provision of the service and opt for content managers that use their own technical *cookies.* Likewise, in the case of mobile applications, analogous precautions need to be applied[17], mainly with

[11] See techniques such as Cookie Syncing or CNAME Cloaking in the risks section.

[12] Wikipedia includes a very exhaustive page on CMS, containing references to the most used CMShttps://en.wikipedia.org/wiki/Content_management_system

[13] Wikipedia has an article in English that also addresses this subjecthttps://en.wikipedia.org/wiki/Mobile_development_framework

[14] An SDK, or Software Development Kit is a set of tools of software development that allows a developer to create an application for a specific system. They usually include an application interface (or API), an integrated development environment (or IDE) and other utilities, included example codes, technical notes and supporting documents to help the programmer develop their task.

[15] User control in the customisation of adverts: https://www.aepd.es/sites/default/files/2019-12/nota-tecnica-android-advertising-id.pdf

[17] The duty to inform and other proactive responsibility measures in apps for mobile devices https://www.aepd.es/sites/default/files/2019-11/nota-tecnica-apps-moviles.pdf

regard to the tracking techniques that are most used in such an environment, in order to avoid for illegitimate processing activities regarding personal data to be added through third-party SDKs. Apart from establishing it as a requirement for the procurement and request it from their providers, this behaviour should likewise be audited by the controller to verify that there is compliance with it.

On another note, in no event will the reject of the use of unnecessary *cookies* for the implementation of a service be a hindrance for access to the portal of the Public Administration, and no less for a limitation or prevention of access to rights and freedoms. In this regard, it must be taken into account that the use of certain technologies, as a result of their specific requirements, may make browsing more difficult as well as access to relevant contents to users who freely decide to not accept *cookies* or who, due to the technological equipment available for them or the limitations thereof (many users use equipment that is not updated, mobile phones with a limited functionality or use accessibility aids or adapted devices) do not support the use of such technologies. In this context, it is important to take into account, at the time to design both the portals or the mobile applications of the public sectors and the contents that are served therein, that the Public Administrations re obliged to meet Royal Decree 1112/2018 of 7 September on accessibility to websites and applications for mobile phones of the public sector [18] [19].

In any event, if we chose to opt for this technology in order to perform an analysis of the traffic, either partially to measure, for example, the interest of users in certain sections of the web portal, it is convenient to follow the data minimisation principle and to collect the browsing data either anonymously or through a pseudonym, and to establish measures [20] that allow to lose the connection with any personal information or, at the very least, to minimise it.

Last, if a third-party tool is used to perform an analysis on the traffic and collect information on the frequency of the visits of the users or the time they spend browsing the different alternatives offered by the market need to be considered, and their technical characteristics need to be analysed, together with the functioning mode[21] and setting options in order to avoid breaching any legal requirement.

Use of tracking technologies in third-party components

As exposed in the section above, portals and mobile applications usually are a great container or structure where elements provided by third-party suppliers are loaded. By browsing through the different Internet pages, several components are indeed being accessed such as images, embedded videos, letter fonts, styles, adverts, statistical counters[22], widgets of social networks, or different *plugins* that are distributed into several third-party servers, and each of them can install its own *cookie* or count our visit. When Public Administrations build their services using such components, they must audit whether such

[18] Royal Decree 1112/2018 of 7 September on accessibility to websites and applications for mobile phones of the public sector https://www.boe.es/diario_boe/txt.php?id=BOE-A-2018-12699

[20] Article 5.1.c of the GDRP: Principles relating to processing of personal data Minimisation Principle

[21] For example, the tool Google Analytics is specially problematic regarding the data protection standards, as the IP addresses of users are stored in serves that are located in the United States, and the recent judgement issued by the European Court of Justice has declared as invalid the "Privacy Shield" agreement that existed between the EU and the USA. (Decision 2016/1250 on the adequacy of the protection provided by the EU-U.S. Privacy Shield.

[22] Cookies and user identification https://developers.google.com/analytics/devguides/collection/analyticsjs/cookies-user-id?hl=es-419

components make use of tracking technologies, not only when the system is being designed, but also during the operation period (life cycle of the system and the processing) and be aware in case tracking elements appear in the updates of their official portals, but also in electronic email messages with HTML, [23]or images[24], news threads RSS[25]or any other product[26].

C. RISKS FOR THE RIGHTS AND FREEDOMS ASSOCIATED TO THE USE OF COOKIES

One of the main risks in the use of *cookies* or other tracking technologies is the collection of personal information beyond what is necessary for the purpose of the processing, especially with regard to special data categories, either directly or indirectly, or by providing the means so that third-parties may do so. Therefore, it is necessary to assess, even for technical *cookies*, to what extent their use may entail a collection or inference of additional information on the subject and to take the necessary measures to minimise such risk.

This circumstance will be more serious when tracking technologies are used by third parties, such as in the case of tracking technologies used by components that are embedded within the web page or a mobile application that pertains to the Administration. This may occur even indirectly when the webpages or apps are built using elements such as letter fonts provided by third parties, images, maps, tools or other applications. Also, when, in order to facilitate the availability of the page, the content is distributed into multiple servers that do not belong to the controller.

The risks in the use of tools for the implementation of websites, such as the referred CMS, must be taken into account, as they usually are updated dynamically. The settings may vary from one version to another, as a result of automated updates, and, in some cases, such updates may not be subject to a prior quality control, and therefore, they may include new cookies or third-party components that carry out additional processing activities. Therefore, the updates or changes must be documented, and subject to prior verifications to their launching, and, by way of a precaution, they must be limited to those new versions with contrasted improvements in relevant subjects for the processing, such as security improvements.

One element to manage such a risk, although not exclusive, is performing regular audits on the components, the type of data that is being processed, the destination of the traffic generated by the applications, the degree of connection of such information with the user's operations in the Public Administrations, and an analysis of the additional inferences and collateral effects that could be performed on the individual subject to administration.

For the event that the Public Administration provides services where the need does not exist to identify the user, or when it is even advisable to have an absence of identification, such as in consulting services for minors or victims, health queries, whistleblowing, etc., the

[23] HTML means Hyper Text Markup Language, the language of webpages. https://en.wikipedia.org/wiki/HTML

[24] RSS means Really Simple Syndication, a news aggregation protocol, whose use has been notably reduced these past years due to the presence of the network called Twitter as a means of rebroadcasting https://en.wikipedia.org/wiki/RSS

risk must be analysed of reidentification, profiling, or register of the historic of the browsing that entails the use of *cookies* and other tracking technologies.

Techniques known as o Cookie Syncing[27] and CNAME Cloaking[28] allow to disguise third-party cookies as first-party cookies in the user's browser, which entails a direct risk for personal data protection by allowing for the controls established in browsers to limit third-party cookies or tracking cookies to be bypassed.

Finally, there are certain security risks regarding cookies that may affect the protection of personal data such as the capture of authentication information in channels that are not secure and log-in theft, *cross-site scripting* (XSS) techniques, use of malicious cookies injected by subdomains (*cookie tossing*), etc. Attacks of this type evolve in time and involve the necessity to reassess the robustness of the web portal according to the state of the art.

[27] What is Cookie Syncing and How Does it Work? (https://clearcode.cc/blog/cookie-syncing/)

[28] Characterizing CNAME cloaking-based tracking (https://blog.apnic.net/2020/08/04/characterizing-cname-cloaking-based-tracking/)

III. SOCIAL NETWORKS

A. WHAT ARE SOCIAL NETWORKS

Social networks are one of the information channels more widely used on the Internet. In fact, together with television (72%), webs and newspaper applications (44%) social networks (53%) are the means that are most frequently used by internet users to obtain information[29].

In social networks, the information contained could be of several types:

- Publications of documents, texts, photographs, videos, links, physical activity data, etc.
- Comments and answers to the publications, creating a thread of dialogue between publishers and readers.
- User profiles, such users being both individuals and organisations. They contain personal information and contact information, as well as the historic of the publications and with what other elements of the network they are related ([30]friends, followers, etc.). Social networks offer different protection mechanisms of the users' contents.

Users find an easy, accessible and immediate channel to share content, socialise or even show a shopping window to offer products or services through social networks, normally without a cost (or even receiving money for their publications, as is the case with *influencers*[31]). A processing occurs in social networks that extends further from the mere exchange of contents between the users and that allows to monetise (to earn money) the activity of the social network. On the one side, the social network profiles the users to publish customised adverts, which have a great impact for the fact of being adapted to the user's tastes and characteristics. On the other side, this social network could sell user data to companies or institutions with an interest in profiling their clients or their prospects in order to offer products or services to them through other channels. These processing activities are usually legitimised, either through the user's consent or within the frame of the execution of a contract, either with or without an economic consideration

In order to be able to make publications, and, sometimes, in order to access content in a social network, it is compulsory to register as a user. Depending on the social network, it may be compulsory for the user to identify with a real identity, thus avoiding that the user may register anonymously, with a pseudonym[32], or that the user may hold several accounts with several identities. Other networks, however, preserve anonymity for their users, at least at the time to register.

Within social networks, the standards on acceptable publications or etiquette are established by the provided or the service, who generally supports a community or a participating structure[33]. Social networks, in the case of European users, are subject to EU legislation, such as the GDPR, and all current legislation where their users are located. This multiple jurisdiction sometimes creates tensions between providers and national authorities,

[29]Digital News Report de 2019 (http://www.digitalnewsreport.es/2019/el-45-de-los-usuarios-elige-la-television-como-medio-principal-para-informarse-mientras-el-40-opta-por-las-fuentes-online/)

[30] The AEPD Blog contains an article with the title ¿Cuánto sabe Facebook sobre mí?- How much does Facebook know about me?https://www.aepd.es/es/prensa-y-comunicacion/blog/cuanto-sabe-facebook-sobre-mi

[31] This term does not exist in Spanish yet. Its definition in English provided by https://dictionary.cambridge.org/es/diccionario/ingles/influencer :*someone who affects or changes the way that other people behave/ a person who is paid by a company to show and describe its products and services on social media, encouraging other people to buy them*

[32] https://www.cnet.com/news/facebook-took-down-more-than-3-billion-fake-accounts/ In May 2019 Facebook estimated that 5% of its active accounts were fake. Twitter erased 70 million fake accounts in 2018 (https://www.bbc.com/news/technology-44682354)

[33] By way of an example we could mention Facebook Community Standards (https://www.facebook.com/communitystandards/)

like, for example, at the time to eliminate content on a specific person[34], fake news, or contents that go against the rights of minorities or especially vulnerable groups.

Regarding personal data processed in interaction with the social network, it must be taken into account that it is not only information associated to the profile of the user that is accessing the network, but also the information generated during the interaction with the content that has been published (which, depending on the type of social network, can be “Like”, reaction-to-content buttons, classification tags, time seals, forwards or republishing, comments, etc.) of general browsing information that include the possibility of data collection such as: webs visited, IP addresses or other localisation information, *fingerprint* of the device, of the browser, an cookies.

The interaction with the social network may also be carried out through a third-party site through the inclusion of widgets of social networks. These widgets may be icons that redirect to the social network. Behind the widgets there are cookies and other information processing mechanisms of browsing, such as what has been explained above, as well as cookies and information processing mechanisms related to the specific interaction in the third-party page where the widget is located.

B. THE SOCIAL NETWORKS AND THE PUBLIC ADMINISTRATIONS

General Aspects

Contrarily to the web sites that are directly or indirectly managed by the Public Administrations themselves and over which a full control of the content or the type of service provided is possible, the social network is an environment that is not envisaged for administrative use, a priori, and that is not adapted to it in general terms. In any event, it entails that the processing be performed pursuant to the provisions in the GDPR.

On certain occasions, social networks are used as a vehicle to reproduce official information (republication or link) and, at the same time, to provide swift information to users on how it would be done, for instance, through a phone channel or an information window channel[35]. In this scenario, the role played by *community managers* or social network administrators is very important. They are those in charge of generating interesting content for citizens with regard to the jurisdiction of the Administration, and they are in charge of providing an answer to questions received through this channel. The role of *community managers* always must be framed within the objectives defined by the body, as part of a communication strategy[36], and they must be aware of the obligation to apply data protection regulations.

The relation of the Public Administrations and the social networks can take many shapes. The most common would be the presence of “official” profiles of the bodies in open or commercial social networks, although other uses cannot be neglected, such as closed social

[34] When we speak about the right to eliminate contents affecting a specific person, we refer to the right to be forgotten. The AEPD has a page with more details on the exercise of this right at https://www.aepd.es/es/areas-de-actuacion/internet-y-redes-sociales/derecho-al-olvido

[35] Several ministries and Public Administrations have their own Facebook account, such as the Ministry of Health: https://www.facebook.com/MinSanidad/. The very Guía de Comunicación Digital para la Administración General del Estado-Digital Communication Guide for the General Administration of the State (https://administracionelectronica.gob.es/pae_Home/pae_Documentacion/pae_Metodolog/pae_Guia_de_Comunicacion_Digital_para_la_Administracion_General_del_Estado.html) in issue 8 thereof provides guidelines for Administrators in social networks. On another note, the Guía de redes sociales de la Generalitat de Cataluña-Generalitat of Catalonia Guide on social networks (http://atenciociutadana.gencat.cat/ca/serveis/xarxes-i-missatgeria-instantania/xarxes-socials/guies-i-normativa/guia-de-xarxes-socials/) is another useful resource for Administrations.

[36] The Guía de Comunicación Digital para la Administración General del Estado provides a frame of criteria, recommendations and good practices to be taken into account by its Departments and bodies at the time to create and generate contents or at the time to update the sites or web portals or the sites related to the new web technologies 2.0 (blogs, accounts or profiles in social networks accessed under the official names of the General Administration of the State bodies or Departments).

networks used as an internal communication vehicle, or networks or groups of employees that are not official, or employees publishing on social networks as individuals, and, last, we could also mention the appearance and reference to the Administrations in third-party profiles, publications or comments.

Internal use of social networks

Administrations may use social networks internally or to connect with the general public. As part of the internal use, we may find close networks, envisaged for employees as an evolution of the corporate intranets with the addition of certain social features[37]. These networks are supported by specific software packages and access is restricted or even forbidden for third parties.

Another possible internal use of social networks is that arising out of informal communications among employees, who create groups to chat or to exchange messages and news as they would normally do in the corridors of the organisation. This type of networks usually appears spontaneously and other times, it can be promoted by the organisations themselves so as to create a bond among their employees. We can find employees pertaining to the same organisation (and, sometimes, together with external individuals and former employees), who talk to one another through a network that is not secure and, sometimes, without being aware of it. An employee may, unconsciously, and without assessing the risks attached, mix the social network where they chat with their colleagues with the Intranet or the internal channels established and publish documents, strategies, or corporate data that may be accessible and even easily copied, by unauthorised third parties. At the same time, this employee may also publish opinions or make eminently personal comments that any given malicious reader may try to present as official statements. It is within this very context and in deep connection with the above, that the AEPD has insisted on several occasions on the dangers that employee groups and comments made in them may present, as they may materialise into harassment at the workplace, sexual harassment and discrimination cases.

The limits of the use of social networks outside the strictly personal environment need to be clearly defined in the Information Policy of the entity, as well as the use of the BYOD *(Bring Your Own Device)*. If these types of communication channels are allowed as part of the policy, a processing is being admitted of which the entity will be controller, and, therefore, such processing will need to comply with the provisions in the GDPR and the ENS (Additional Provision I of the LOPDGDD).

Employees who publish on their social network as individuals

Social networks had provided the possibility to many people of sharing their knowledge or ideas either anonymously or with their own name. These people have created some sort of personal brand under which prestige is measured according to the number of followers, links and mentions. As part of the more successful individuals mentioned above, we find the *influencers*, who have built their fame on the publication of their opinions in social networks.

Without reaching this level, it is frequent for professionals of all sectors to publish contents in social networks adding their experience to common knowledge. In some companies, the use of social networks is even encouraged by experts in the organisation as a means of promotion[38], especially in social networks where their intention is to share the employee's work experience. Notwithstanding, the fact of employees publishing professional information

[37] Some examples are IBM Notes, Jive or Ms Sharepoint

[38] 5 ways to turn employees into advocates of your brand on Twitter https://business.twitter.com/es/blog/employees-advocates-on-twitter.html

may breach the data protection principles when personal information is included and there is an absence of legitimisation.

The Information Policy of the entity needs to clearly state, at least from the perspective of data protection, the prohibition of processing personal data regarding the entity's activity through the personal accounts of employees in social networks, and that employees who engage in such a behaviour will otherwise be considered as data controllers.

The use of social networks by Public Administrations in order to offer services to their citizens

When a Public Administration decides to offer information or services to citizens through a social network, the citizen that is subject to administration cannot be obliged to have an account in such a network inasmuch as such networks perform additional processing activities based on the user's consent.

Pursuant to Recitals 42 and 43 and Article 7 of the GDPR for a processing based on the user's consent to be licit, such a consent needs to comply with a series of conditions, more precisely, such a consent must be informed, specific, freely given, and unequivocal, and it must further be granted in conditions where there is not an asymmetry between the data subject and the data controller. Should the Administration provide this channel as the sole means for a service without providing any further alternative channels through which such service can be provided with a field level interaction, such a consent may not be considered as freely given, and it will further be considered as an obstacle [39]for such citizens that are most affected by the digital breach in the face of the impossibility to access the information provided or to exercise the rights they are entitled to exert.

For example, the fact of performing citizen involvement initiatives solely through social networks forces any data subject with a wish to influence such an initiative to consent to the processing by a third party and, consequently, to grant a consent that would not comply with the requirements envisaged in the GDPR.

The service provided through a social network must comply with all obligations provided for in the GDPR, among which, the duty to inform. This obligation could be implemented through a post fixed at the top of the account in such a way that the user may easily access such a post and that a privacy policy is added to such post or a link to such policy.

The implementation of services in social networks that are specifically oriented towards minors is a special case. In these situations, when the activity in the social network implies the processing of personal data of minors under 14, it must be guaranteed that the consent for the processing has been granted by their parents or guardians.

The processing of personal data by networks where users have an account will generally be based on the data needed for the execution of the service contract or on the consent provided. The difference must be made between such grounds and the legitimacy of data processing activities of the citizens' data performed by the public administrations with an account in social networks arising out of the interaction with people as a result of the information provided by the Administration through their official account and whose legitimacy to be processed by the Public Administration would be based on the fulfilment of a mission carried out for the public interest or through the exercise of the public powers granted to them.

[39] Article 14 of Act 39/2015, of 1 October, on the common administrative procedure, establishes the rights for individuals to choose at all times whether they wish to communicate with Public Administrations for the exercise of their rights and obligations through electronic means, except as otherwise provided by the applicable regulations. https://www.boe.es/buscar/act.php?id=BOE-A-2015-10565&p=20200911&tn=1#a14

Furthermore, vis-à-vis the responsibilities undertaken by the Public Administration, and regarding the tracking technologies (cookies) covered above in this document, the processing of data performed by the social network could involve the collection of a greater volume of personal data pertaining to the citizen that exceeds what is needed by the Public Administration in the management of their official account or of additional data that have no relation whatsoever with such a management. The role played by the Public Administration and the social network with regard to the personal data processing could change to that of joint controllers depending on whether such Public Administration is not diligent to know of such a circumstance or, if the Public Administration is aware of them, it is within its power to configure or to limit the processing performed by the social network and does not use such data, or actively allows the processing activity[40]. By way of an example, in the case of widgets to social networks included in the Public Administration portals, before allowing the use of cookies by such resources, Public Administrations need to obtain the visitor's prior consent of their pages, while keeping the widgets disabled until consent has been obtained. The Public Administrations, at all times, have to guarantee that the data processed are strictly necessary for the fulfilment of the legitimate interests by way of compliance of the minimisation principle.

The majority of the social networks do not offer sufficiently contrasted quality levels of service so that they can become notice instruments, as the integrity of the information provided cannot be guaranteed and that such information is not going to be shortened or enriched with advertising or other third-party elements. Neither are there confidentiality guarantees in the transfer of information to a social network, as in some of them it is possible to access content by the very social network [41], or else the privacy setting of the user may grant an undiscriminating access to the publications made, a consent that is not implicit on the content submitted by an Administration.

Usually, the entity must keep a policy on communication on social networks that must reflect aspects of the Data Protection Policy[42] of the entity, such as:

- A clear responsibility efficiently embedded within the responsibility change of the organisation, in such a way that the communication through the network or a social network is in line with the rest of the institutional communication policy.
- An information document for public employees where instructions are given on what to do and not to do on the social network of the Administration. If we start from the premise that social networks are a support or a complement to the web page and an official site, it is easier to limit the service and to avoid risks.
- An adequate training of the persons in charge of publishing the contents and meeting the comments and the questions, detailing:
  - Which personal information may be published and which may not be published.
  - In the case that documents are published that have been abstracted or anonymised, this must be performed with sufficient guarantees so as not to

[40] Pursuant to the provisions by the EDPB in their document named https://edpb.europa.eu/sites/edpb/files/consultation/edpb_guidelines_202007_controllerprocessor_en.pdfGuidelines 07/2020 with regard to the definition of data controller and data processor (https://edpb.europa.eu/sites/edpb/files/consultation/edpb_guidelines_202007_controllerprocessor_en.pdf) and the Judgement by the European Court of Justice Unabhängiges Landeszentrum für Datenschutz Schleswig-Holstein v Wirtschaftsakademie Schleswig-Holstein GmbH, Case C-210/16, 5 June 2018, it can entail the joint control of the processing of the data collected. Notwithstanding this last judgement rules that the fact of using a social network does not imply for the role of joint controller to appear, but rather, when there is involvement in the establishment of the purpose and the means of the processing.

[41] Some social networks access the contents to profile the user for marketing purposes. Other may entail a breach of privacy through the application of policies against child abuse, digital violence, psychological analysis, etc., in case of a wrong construction of official communications.

[42] Article 24 of the GDRP:Responsibility of the controller

reveal, for example, the data of the electronic signature of the signing party or the secure verification code [43] that would allow to recover all the document.

- The way in which they should be dialogue with the users.
- How do the messages or notices containing personal data and requesting attention that have been received through the channel need to be escalated or internally communicated, such as requests of removal of content, requests revealing the data of the requesting party to third parties or terrorist threats or suicidal messages.

Social networks may be a means for the Administrations to know statistical data about[44] their users, through the information provided by the platform itself. This could lead to a profiling of the users, their interests and their browsing habits. To the extent that they involve personal data processing, apart from being legitimised, especially when special data categories may be inferred from the browsing, information needs to be provided on such processing activities pursuant to the provisions in Articles 13and 14 of the GDPR.

It must be kept in mind that some social networks allow to forward content to third parties. If content may exist that includes personal data, Recital 66 of the GDPR will apply. *To strengthen the right to be forgotten in the online environment, the right to erasure should also be extended in such a way that a controller who has made the personal data public should be obliged to inform the controllers which are processing such personal data to erase any links to, or copies or replications of those personal data. In doing so, that controller should take reasonable steps, taking into account available technology and the means available to the controller, including technical measures, to inform the controllers which are processing the personal data of the data subject's request.*

The fact of searching a positioning or an influence, as can be performed in other backgrounds, even resorting to illicit techniques such as the purchase of followers[45], the publication of fake opinions, or the abuse of tags and of the most searched terms, could be construed as a lack of transparency and lawfulness with regard to the processing that is based on the use of a social network without losing sight of the absence of data quality[46] linked to the breach of the accuracy principle

C. RISKS FOR THE RIGHTS AND FREEDOMS ASSOCIATED TO SOCIAL NETWORKS OF THE ADMINISTRATIONS

Once that a processing has been designed through the social networks that complies with the requirements of the GDPR, it is necessary to analyse the risks for the rights and freedoms of the citizens that may arise for the purpose of performing an adequate management thereof.

One of the risks is the occurrence of errors in the application of communication policies through social networks that reveal personal data unwillingly either of the individuals subject to administration or of the public employees themselves (publication of CSV[47], metadata, etc.). In order to reduce the risk, it is necessary to define an internal policy that is clear and broadly known by the personnel on the implications and the consequences of being involved in social networks.

[43] Secure verification code in the Spanish Government: https://sede.administracion.gob.es/PAG_Sede/ayuda/ayudaCSV.html

[44] For example, Facebook offers statistics for professional pages or *fan pages,* contrarily to personal profiles or pages. (https://es-la.facebook.com/business/insights/tools/audience-insights)

[45] Article on the purchase of fake followers on social networks https://www.genbeta.com/redes-sociales-y-comunidades/hemos-comprado-seguidores-falsos-twitter-instagram-facebook-se-nos-ha-quedado-cara-tonto

[46] Article 5.1.d of the GDRP:Principles relating to processing of personal data Accuracy principle

[47] https://www.rtve.es/noticias/20181018/cgpj-atribuye-fallo-sistemico-filtracion-datos-victima-manada/1822280.shtml

In this policy both the guidelines of the management of official accounts in internal networks and the advice on involvement, as an individual, in unofficial networks need to be defined. More precisely, they must specify which content can be published in social networks and which content must not[48], together with the consequences for those breaking these rules, training activities to be taken by the employees, a regular audit of the publication processes, and a permanent follow-up, by the persons of the organisation in charge of managing the entity's presence on social networks, so as to ensure the security of the processing activities performed within this context and to provide a proactive answer to any incident[49].

For example, the distribution of the content by the Public Administrations through the services of instant messaging, the use of distribution lists or distribution channels, instead of groups, avoids the identification of the members among each other.

Apart from training involved personnel, it is likewise advisable to verify on a regular basis that the politics defined by the entity regarding internal social networks are followed, as well as the use of BYOD, or the use of personal networks where personal information regarding the entity is published.

Among the elements that need to be verified and continuously monitored we could mention:

- Personal data of the employees of the Administration. Do individuals make publications in their own name, or are full ID data, wages, or assets of executive members directly revealed on the network? Are photographs of the premises or internal events published without the adequate measures being adopted?
- Files and cases processed by the Administration: Are examples or full resolutions published that have been badly anonymised, or is information provided on the status of the procedure to individuals that claim to be data subjects without making any kind of prior verification?
- Applicants that seek to be employees or students attending training courses: Are lists, marks, or photographs taken during the course published without due precaution?
- Users of a social network: Are data of their identity revealed in the comments?
- Data inferred that may be deducted regarding a natural person determined as a result of the publication.

Tools to make this verification may be, for example, the regular review of the contents, either manually or automatically, the performance of surveys among employees or regular audits.

With regard to the publication of anonymised information, as stated above, the anonymised content that has been published needs to be subject to review both automatically and manually for the purposes of guaranteeing that an error has not occurred during the anonymisation procedure that exposes personal data of citizens subject to administration or public employees. Certain procedures need to be included in order to verify the metadata that may be included within the documents that are being published, both those revealing information on public employees and other data that allow to access the original information of the anonymised documents, such as CSVs.

In anonymisation processes of the content that is to be published on a social network, it is advisable that a clear compartmentalisation be performed between the anonymised content and the content that has not been anonymised, as well as other guarantees that hinder the

[48] A more precise term is Social Media Policy. https://www.forbes.com/sites/forbeshumanresourcescouncil/2017/05/25/why-your-business-needs-a-social-media-policy-and-eight-things-it-should-cover/#3f09e7495264

[49] Article 32 of the GDRP:Security of processing

possibility of a publication of contents with personal data by mistake. More precisely, it is advisable to avoid that the manager of the social network have access to the non-anonymised content easily.

Pursuant to the minimisation principle, whenever identification of the citizen subject to administration is not compulsory for the processing of the identification, it is advisable to avoid the collection and the processing of such a datum. There is a risk of reidentification and profiling of the citizens subject to administration through the analysis of the visits performed as well as through the visits to specific content. Such as with cookies, there are more sensitive services, such as counselling for minors or victims, health counselling, whistleblowing mail boxes, etc., that may compromise the user themselves as browsing habits are recorded, content is shown based on the context profiling or the referred reidentification, all of the generating risks that need to be managed.

For example, whenever processes are being conducted of involvement or survey of the citizens through a social network, the risks must be analysed of reidentification by third parties linked to the metadata collected during the survey process, such as IP addresses or signatures from the devices that are being utilised by the users, as well as the possible link of then answers with personal data held by the Public Administration.

Accounts opened in a social network by a Public Administration can be compromised and, therefore, content can be leaked in such accounts for the purposes of collecting personal data. This must be taken into account at the time to manage the processes of assignment and renewal of access codes.

On another note, third-party content, inserted in the shape of adverts or otherwise, may lead users to believe that they are links that pertain to the services of the Public Administration that collect personal data. Therefore, such content needs to be supervised. Furthermore, the type of content that can be offered needs to be controlled and, more precisely, the content that relates to ideologies and beliefs of any nature.

In the event that spaces are provided to the citizens subject to administration to upload their own content, it is advisable to avoid that such publication of the referred contents is made unsupervised so as to avoid the dissemination of third-party personal information without their consent, more precisely, sensitive content or any behaviour that would imply acts of digital violence.

In case that access to the content of the social network requires a log in by the citizen subject to administration, the privacy guarantees need to be assessed by the Public Administration, including security measures oriented towards privacy and provided by such network, namely the user creation and password creation policy. This analysis must be performed for the sake of citizens, but also for the sake of the managers of this space, as the attacker could impersonate the publishers[50].

In addition, vis-à-vis the responsibilities undertaken by the Public Administrations, and regarding the tracking technologies mentioned in the section above, it is preferable that static links to the official account in the social network be included in the Administration site or the *newsletter* that are to be distributed, and refrain from using *widgets* that are embedded within the content.

In this case, the Public Administrations are under the obligation of verifying that such link directs citizens to a service on the Internet that is compliant with the regulation. For example, if such website uses cookies of which information needs to be provided and prior consent is requested for processing activities such as profiling and tracking, it is necessary for a regular

[50] A post can be found at the AEPD blog that addresses Security measures on Facebook (https://www.aepd.es/es/prensa-y-comunicacion/blog/medidas-de-seguridad-en-facebook)

supervision to be in place so that it can be verified that such circumstance is being met efficiently.

Finally, even though the subject of compliance obligations has been addressed before regarding minors, it must always be kept in mind that current techniques do not always allow to establish that the user is a minor with full effectiveness. The risk will need to be assessed that exists with regard to the type of contents that are being offered through the network and, according to this risk, additional procedures will need to be implemented for a more precise verification of the users' age.

IV. CLOUD COMPUTING

A. What is cloud computing?

Cloud computing[51] is a way to use servers in remote locations in a flexible and transparent manner. Instead of having their own equipment, either bought or hosted in the facilities of an organisation, a provider "rents" virtual equipment or specific services that are managed through the network. This model suppresses the problems of initial investments and obsolescence, thus allowing the client to pay for the computing capacity, the bandwidth, or the services they need at each moment. The supplier of the service may guarantee the parameters of the quality of the service by means of a contract, as well as the privacy guarantees, the security guarantees regarding security oriented towards other purposes, maintenance, software packages updates, etc.

Although during the first times of cloud computing the tools offered by the provided were equal to those that a client could purchase and store in their machines, today, there are databases, libraries, programming languages, and specific software components for the cloud[52], as well as certificates[53] specialised professionals and a wide range of academic discipline[54].

Depending on the type of services that are being offered to the client, we can find IaaS solutions (Infrastructure as a Service) when the provided solely provides storage capacity and mass processing; PaaS (Platform as a Service), if the basic utilities are provided in order to build applications on which solutions are developed, and SaaS (Software as a Service) when the client finds all the final tools in the cloud needed to implement the processes of their organisation. Another possible classification of the clouds is the separation between public and private clouds. A public cloud is when providers provide "independent" services to different clients, such as space for websites, storage or process capacity. On the other hand, we talk about a private cloud when the provider offers a series of closed services in a package to third parties. We could compare a private cloud to a private network or the intranet of an organisation, with its elements isolated from the outside, excess for certain localised accesses, and a centralised management.

With the popularisation of the services in the cloud, services were sometimes offered that were not transparent about the location where data were being processed[55]. This virtualisation and relocation are not without risks both from the point of view of security and from the point of view of personal data protection[56].

B. Cloud computing and the Public Administrations

The Public Administrations use the cloud as part of the services provided to citizens and as an element of internal management. In fact, the General Secretariat of Digital Administration provides a hybrid cloud computation and storage service to the General Administration of the State and its public bodies through the shared service Nube SARA[57].

[51] Wikipedia https://en.wikipedia.org/wiki/Cloud_computing

[52] Amazon web Services includes up to 21 specific product categories for its cloud. https://aws.amazon.com/es/products/

[53] The most popular ones are AWS, CISCO, Microsoft and WMWARE

[54] Google Scholar delivers 6310 papers on Cloud computing only in 2020. https://scholar.google.es/scholar?as_ylo=2020&q=academic+papers+cloud+computing&hl=es&as_sdt=0,5&as_vis=1

[55]Azure regions https://azure.microsoft.com/es-es/global-infrastructure/regions/

[56] The AEPD has a Guide for clients who hire Cloud Computing services (https://www.aepd.es/sites/default/files/2019-09/guia-cloud-clientes.pdf) and Guidelines for Service Providers of Cloud Computing (https://www.aepd.es/sites/default/files/2019-09/guia-cloud-prestadores.pdf).

[57] According to the Digital Administration service catalogue published on the Electronic Administration Portal, Nube SARA is part of the services declared as shared services (CETIC 15/09/2015) and provides computing and storage services in a hybrid cloud for the GAE and

Services in the cloud allow for Administrations to have a "generic brand" and adapt them and deploy them[58] without an asset investment opting for a more flexible and scalable governance model, which could be interesting in the case of small Administrations without many resources, such as Local Governments and Local Entities. In order for this facility not to become a problem, it is necessary to carefully plain, at least from the point of view of compliance with data protection, the way to implement these services as well as with which *partners* they are to be implemented.

In addition to the above, the cloud does not permit a direct control on the services, and, when errors, breaches or discontinuities of the service occur, we need to rely on the knowledge of external third-parties and response times regulated by SLA or service level agreements that are established by contract. When the option is chosen to hire services in the cloud, it is the data controller's responsibility to guarantee that the suitable provider is selected[59] and to guarantee that the contract contains the requested clauses and guarantees that bind the service provider in the fulfilment of data protection regulations, as well as any other applicable regulation for a specific public entity. In any event, it is important to take into account that the advertising information or the descriptive technical documents published in the provider0s webpage on the cloud services offered are not, on their own, contractual and binding terms with regard to the service provision, unless they are added to the offer signed with the assignee provider of the tender. Under no circumstances can a generic standard-form agreement for a specific service be construed as binding when conditions are established apart from the general contracting requirements for Public Administrations established by the Public Section Procurement Act[60].

Often, there is a tendency to identify the cloud with the public cloud, which is cheaper and more popular, leaving aside the possibility if setting and sharing private clouds for an Administration or a set of them. There are providers that offer private clouds with security certificates thus offering their clients not only the ability to develop their services, but also to certify the security of their Administrations.

The procurement of a service in the cloud does not entail the full disappearance of the security obligations of the data processor, but rather, the data controller will always be in charge of any decision-making with regard to the requirements of the personal data protection where, necessarily, the requirements established in Article 32 of the GDPR in terms of security need to be present. Apart from the supervisory obligations that lie with the data controller on the measures of which the data processor is in charge, the data controller will always have the obligation to implement any of such measures, such as the definition and the management of the access control policy.

Another obligation that lies with the Public Administration is that of requesting and obtaining information on whether third-parties are involved or not (subcontractors) in the service provision of the *cloud computing* and to exercise a control thereof as described in the

its Public Bodies through the setting of consolidation nodes both in CPDs of the Administration (private cloud) and of external providers (public cloud), which allows for the clients' TIC units to obtain capacities both in the private cloud and in the public cloud for each of the services that they need to implement depending on their features and the costs they can undertake. Even though it initially focuses on infrastructure as a service, the idea is to gradually provide more mature services, such as platform as a service or application as a service, for example, the management of the payroll in the cloud). The components of the service that are provided will be the usual in services of these characteristics: virtual servers provision with different customised features, shared storage, communications, backup, high component availability, monitoring, consumption control, ...

[58] There are several technologies that permit the deployment of instances of services swiftly, in a faster way than virtual machines. To learn more, visit, for instance, Docker (https://www.docker.com/resources/what-container) or Kubernetes (https://kubernetes.io/es/docs/concepts/overview/what-is-kubernetes/)

Article 28.1 GDPR: "*Where processing is to be carried out on behalf of a controller, the controller shall use only processors providing sufficient guarantees to implement appropriate technical and organisational measures in such a manner that processing will meet the requirements of this Regulation and ensure the protection of the rights of the data subject*".

[60] https://www.boe.es/buscar/act.php?id=BOE-A-2017-12902

Guide for clients hiring Computing services[61]. In addition to the above, the physical location of the servers that are to store the personal data needs to be clarified in the cloud contract. If the servers are located in another country, we need to ensure that the equivalent guarantees in terms of personal data protection are kept.

There is another restriction regarding the location of the provider's servers for *cloud computing* services that has been hired introduced by Royal Decree Law 14/2019 of 31 October, adopting urgent measures by reason of public security in terms of Digital Administration, public procurement, and telecommunications[62]. This regulation, through the amendment of Act 9/2017 of 8 November on Public Procurement[63], establishes the obligation for the assignee company to submit, before the execution of the contract, a statement where an indication is made on where the servers are to be located and the provision of the services that is associated thereto is to be delivered. This information may be of importance for the sake of national security. It is also necessary to ensure compliance by such contract with national regulations and European regulations in terms of data protection.

The Public Administration needs to take into account the data protection requirements throughout the life cycle of the service provision, not only since the moment of election of the service provider, or during the service provision itself, but rather, once the contractual relation is terminated, it needs to ensure that the information is destroyed, that the contracting entity receives such information, or that the new designated data processor becomes in charge of such information pursuant to the National Scheme of Interoperability[64]. To that end, and for the purposes of ensuring legal compliance, it is advisable to carry out audits and to request regular compliance certifications.

C. RISKS FOR THE RIGHTS AND FREEDOMS ASSOCIATED TO CLOUD COMPUTING

Despite the advantages provided by the cloud, this type of solutions also entails a series of risks that need to be taken into account. Among these risks, there is the privacy of the information stored, as well as the continuity of the services, the legal changes, and the loss of control of the infrastructure and the applications used. In the particular case of Public Administrations, due to the volume and the sensitivity of data to be managed, such risks need to be subject to a thorough analysis for each scenario where the use of this type of solutions is envisaged, more precisely, when the rights and freedoms of the citizens depend on the services deployed in the cloud.

It is likewise important to assess what type of architecture or *cloud,* public or private, entails a lesser level of risk for the specific processing that needs to be implemented. It must be kept in mind that the impossibility by the data controller to technically fulfil any of the obligations requested by the GDPR must not be envisaged as a risk that needs to be dealt with, but rather as a legal breach. Under no circumstances will the requirements in the GDPR and the LLOPDGDD be substituted with technical and organisational measures.

At the time to design the processing, it is advisable to assess the addition and application of data minimisation mechanisms according to the risk, more precisely, by limiting the extent

[61] Detailed in the chapter "What needs to be known about the procurement of Cloud Computing Services", more precisely in section 4.
[62] BOE number. 266, of 5 November 2019 https://www.boe.es/diario_boe/txt.php?id=BOE-A-2019-15790
[63] Act 9/2017, of 8 November, on Public Procurement implementing the Directives of the European Parliament and of the Council numbers 2014/23/UE and 2014/24/UE, dated 26 February 2014. https://www.boe.es/buscar/act.php?id=BOE-A-2017-12902
[64] https://www.boe.es/buscar/act.php?id=BOE-A-2010-1331

of the data[65], uploading anonymised or pseudonymised data to the cloud, using homomorphic encryption, etc.[66].

The data of public employees can also be exposed. The risk must be assessed that the information on the employees, mainly while telecommuting and/or using BYOD, may be exposed to third parties, and more precisely, metadata that can expose their personal lives. The context in which work activities of public employees can be located does not exclude, in any event, the responsibility by the Public Administrations regarding the personal data protection regulations.

At the time to identify the information that is uploaded to the cloud, sometimes, the question arises on whether this information is personal data and whether the minimisation principle is being met. They can be aggregated data, compressed data, or data on which some technique has been applied in order to disassociate such data from natural persons, but this does not necessarily mean that they have ceased to be personal data if people are behind these data that can be subject to identification. This is a question that needs to be analysed from a formal perspective and the reidentification risks must likewise be assessed rigorously and responsibly, thus avoiding simplistic answers as a way to bypass the liability of processing personal data. Such an approach needs to be part of a more comprehensive approach that assesses the maturity level of the anonymisation processes[67] that are being used by the organisation.

During the life of a system, it is important for it to be adequately monitored with regard to its functioning, and that the implementation of the service be audited, as it will evolve in time. This proactive attitude is, at first, the responsibility [68] of the persons in charge of the controller of the service, but such a responsibility needs to extend throughout the chain of assignments and subcontracts. In case of a breach of the data protection regulation, the controller cannot hide behind the fact that they did not know what their subordinates were doing because, as stated in Article 28 of the GDPR, the controller is responsible both for choosing a data processor that offers enough guarantees and for formally indicating, by means of a contract, the guidelines and the data processor's behaviour at the time to process personal data.

It is not improbable that breaches occur in the cloud services that endanger the availability, the integrity or the confidentiality of personal data with consequences for the rights and freedoms of natural persons. A cyber-attack, a bad functioning of the system or a human error may endanger the data of citizens and public employees. As indicated before, the information security risk management does not solely lie with the service provider, but rather, the data controller is responsible for determining the security measures that need to be requested from the data processor and that need to be contractually reflected.

With regard to this last aspect, in case of a security breach affecting personal data, the data controller needs to implement a series of urgent mechanisms [69].. On the one side, they must face the contingency itself in the way that is less risky for the data and the service in general. Simultaneously, and within a period of 72 hours, they must notify the Data Protection

[65] Recital 78 and Article 25 RGPD.

[66] Homomorphic encryption is a technique that allows to perform operations on the encrypted data and obtain results, also encrypted, equivalent to the operations directly performed on the original information. In the post Encryption and Privacy III: Homomorphic Encryption (https://www.aepd.es/es/prensa-y-comunicacion/blog/cifrado-privacidad-iii-cifrado-homomorfico)published by the Spanish Data Protection Agency more information can be obtained on this data minimisation technique.

[67] Privacy Analytics, "The De-identification Maturity Model", Khaled El Emam, PhD, Waël Hassan, PhD. 2013 https://iapp.org/media/pdf/resource_center/2014-14-05%20Privacy%20Analytics%20The%20De-identification%20Maturity%20Model.pdf

[68] Article 5.2 of the GDRP: Principles relating to processing of personal data Proactive responsibility

[69] The AEPD has published a Guide for the management and notification of security breaches. It can be consulted at https://www.aepd.es/sites/default/files/2019-09/guia-brechas-seguridad.pdf . There are also several articles in its blog (https://www.aepd.es/es/prensa-y-comunicacion/blog) on this subject. The WP29 also published some guidelines in February 2018 (https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612052)

Authority on the information that has been breached[70]. This is not solely a statistic or bureaucratic process. It is aimed at creating a more resilient society by allowing the Supervisory authority and other entities to gain knowledge on the materialisation of certain risks and be able to react in a proactive manner. Likewise, in certain situations, the Supervisory Authority may order the data controller to notify the users whose data have been affected by the security breach so that they stay alert or adopt their own security and protection measures vis-à-vis the incident.

The Administration, as a data controller, must also manage the risks in the event the provider of the cloud unilaterally decides to discontinue the service or change the conditions of the provision, as well as manage the legal risk of any changes in the legislation, or any other changes of any nature that prevent the use of such services. This last scenario is particularly relevant in providers that are not located within the national territory. Therefore, the data controller, as an answer to an analysis of the possible risks,[71] must implement the necessary measures and contingency plans, together with service migration strategies to other services.

[70] Article 33 of the GDPR - Notification of a personal data breach to the supervisory authority

[71] Article 32 of the GDRP: Security of processing

V. BIG DATA OR MASS DATA PROCESSING

A. WHAT IS BIG DATA

Pursuant to the definition provided by ISO, when we speak about *Big Data*[72], or mass data processing, we refer to great datasets that are characterised by their high volume, variety, speed and/or variability, which request a scalable technology for an efficient storage, handling, management and analysis.

Mass processing technologies have evolved considerably in the last years giving way to a wide range of processing activities. Internet has made available to all of us a great amount of data that can be used. The Public Administrations themselves, through the initiative called Iniciativa Aporta[73] and through the portal datos.gob.es, promoted by the Ministry of Economic Affairs, Digital Transformation and the Corporate Public Entity Red.es, promotes the availability and the reuse of public information and the development of advanced services that are based on data.

The mass analysis of data has allowed to obtain real-time information, or almost real, out of information sources and datasets that are distributed all around the world. There are specific software platforms to extract, load and transform (the acronym ELT is preferred over the acronym ETL of the dashboards) data of different origins and to exploit such information. This exploitation can take different shapes that would range from a graphic representation, such as, for example, dashboards, to the construction of new information or profiles, thus regrouping the data differently from how they have been obtained.

From the point of view of the protection of personal data, it is vital to ensure that a legitimacy exists for such mass data processing and, in case special data categories are included, it is necessary to priory forbid any processing of such type of data. These conditions not only need to be complied with for data included in the set of data collected, but also with regard to those that can be inferred from the comparison and connection with the original data.

During the design phase of the processing of *Big Data*, the amount of data that is needed and sufficient with regard to the purpose of the processing needs to be analysed in an objective manner[74], and the minimisation principle needs to be complied with[75]. Furthermore, strategies where the maximum amount of data is collected "just in case" without the prior adoption of selecting criteria must be discarded. This problem can be enhanced in case of massive collection of data supported by sensors within processing contexts such as those performed in *Smart Cities*.

This technology allows for the profiling or the enrichment of the profiles of individuals, a processing that requests a legitimacy and needs to meet certain requirements and conditions, among which, those related to automated individual decision-making[76], and, as the case may be, the performance of an impact assessment for the protection of data[77], and when applicable, the prior consultation[78] to the Supervisory Authority.

[72] ISO/IEC 20546:2019 Information Technology – Big Data – Overview and vocabulary

[73] Iniciativa Aporta https://datos.gob.es/sites/default/files/datosgobes/190522_iniciativa_aporta_-_contexto_y_directrices.pdf is developed within the current legal frame https://datos.gob.es/sites/default/files/datosgobes/190522_iniciativa_aporta_-_contexto_y_directrices.pdf and the platform datos.gob.es is articulated as a meeting point between the Administrations, the companies and the citizens that are part of the ecosystem of open data in Spain.

[74] With regard to the application of the minimisation principle, please read the Guide to Data Protection by Default of the AEPD.

[75] Article 5.1.c of the GDRP: Principles relating to processing of personal data Minimisation Principle

[76] Article 22 of the GDPR- Automated individual decision-making, including profiling

[77] Article 36 of the GDPR – Prior consultation

[78] Article 35 of the GDRP: Data protection impact assessment

If personal data have been inferred, we must handle them as personal data, with all guarantees regarding security[79], rights[80], international transfers[81] and others.

If the same data controller of the mass data processing has been in charge of collecting the personal information directly from the data subjects[82], for example, through apps, forms on the Internet, or through the use of *cookies* or other technologies, they must enquire whether they are allowed to do so and, in case this legitimacy[83] is provided through the users' consent, for such consent to be duly obtained[84] and for transparency of the information[80] to be the highest possible so that users are clearly aware of where, how and to what end their data are going to be used and of the existing means to claim their rights over such data, or to withdraw their consent. In case the data controller in charge of the processing has not directly collected the personal data the information obligations and requirements need to be taken into account as well that are provided for in the regulation[85].

If the processing is carried out by a third party acting as data processor[86], as is the case with other technologies that have already been mentioned in this document, a contract needs to be executed that details how the data processing will be carried out since the data are made available and until such data are restored or destroyed, including the security measures and the additional guarantees that need to be adopted.

B. Big Data and the Public Administrations

Big Data, together with other technologies referred to in this document, such as, for example, cloud computing, or artificial intelligence, is a tool that allows to process and extract value out of big volumes of information generated by the Public Administrations.

On the one side, Administrations release many data in simple formats as part of the open data initiatives[87]. These initiatives, that are increasingly more extended, consider the opening of data as a means of transparency[88] and seek to make data referring to population, transportation, background, health, energy, territory, education, etc., available and reusable, such data being data stored by the Public Administrations in their systems. The aim is to provide information to the citizens as an exercise of transparency that helps build more trust in the body and in the corporate sector, so that they integrate such date within their systems and use them in their own processes, thus contributing to the enhancement of economy and innovation.

On another note, those Administrations with the ability to analyse such big datasets have developed equipment and are comparing different sources of information in order to extract knowledge and apply such a potential provided by mass data analysing to different sectors and scenarios such as health, tourism, research, sustainable development, security or anti-fraud.

Notwithstanding, it must be kept in mind that this mass analysis and the comparison of heterogeneous information sources may also bring about negative consequences from an

[79] Article 32 of the GDRP: Security of processing

[80] Article 12 of the GDPR - Transparent information, communication and modalities for the exercise of the rights of the data subject

[81] Article 44 of the GDRP: General Principle for Transfers

[82] Article 13 of the GDPR - Information to be provided where personal data are collected from the data subject

[83] Article 6 of the GDRP: Lawfulness of processing

[84] Article 7 of the GDPR– Conditions for consent

[85] Article 14 of the GDPR - Information to be provided where personal data are collected from the data subject

[86] Article 28 of the GDRP: Data processor

[87] Open Data Initiative by the Spanish Government https://datos.gob.es/

[88] In Europe, we start with Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003, updated by Directive 2013/37/EU of the European Parliament and of the Council of 26 June. In Spain, this has given rise to Act 37/2007 of 16 November on the reutilisation of the information of the public sector, amended by Act 18/2015 of 9 July. More information is available on the regulations at https://datos.gob.es/sites/default/files/datosgobes/190522_iniciativa_aporta_-_contexto_y_directrices.pdf

ethical point of view, in terms of privacy, mainly in data protection, if a wrong use of the information used was to be made.

As is the case with every processing activity, the principles of lawfulness, fairness[89] and limitation of the processing need to be complied with[90]. For processing activities based on *Big Data,* by virtue of their own nature, it seems relatively easy to arrive into situations where the initial purpose of the processing is diluted when the datum is exploited for secondary purposes, as the general data protection regulations do not impede that personal data be used for different purposes from the purpose for which they were collected. They only state that such purposes must not be inconsistent with the initial purpose. Therefore, for them to be reused in new projects, it is key to perform an analysis of lack of incompatibility where the following considerations reflected in Article 6.4 and Recital 50 of the GDPR are taken into account:

- That a relation exists between the original purpose of the processing and other subsequent purposes.
- That such subsequent processing activities be within the reasonable expectations of data subjects.
- The nature and the sensitivity of the data subject to processing.
- The impact that any subsequent processing may have on data subjects.
- That the adequate protection, technical, and organisational measures be adopted.

It is important to take into account that not all legitimate interests can be applied for the case of processing activities performed by the Public Administrations, more precisely, the limitations established for the legitimate interest to apply.

C. RISKS FOR THE RIGHTS AND FREEDOMS ASSOCIATED TO BIG DATA

The General Data Protection Regulation addresses the subject of large scale processing activities and requests for special precautions to be adopted for this type of processing[91] as, from the point of view of personal data protection, the processing of great volumes of information may entail risks that need to be managed[92.] The mass processing of personal data is one of the cases for which the GDPR establishes the requirement of performing a more systematic risk assessment, requesting the performance of an impact assessment regarding data protection and, as the case may be, and according to the result obtained, a prior consultation to the Supervisory Authority.

In this regard, it must be taken into account that the mass character of a processing or the great scale processing of data is not exclusively defined by the amount of data or the number of individuals whose data may be processed at a certain time, pursuant to what is indicated in the document of the Article 29 Working Party of the referred Directive on Guidelines on the Data Protection Officer (DPO), the term “great scale” associated to *Big Data* processes needs to be understood in several dimensions:

- With regard to the number of affected data subjects
- With regard to the volume or the variety of elements of the data that are subject to processing

[89] Article 5.1.a) of the GDPR- Principles relating to processing of personal data Principle of lawfulness, fairness and transparency

[90] Article 5.1.b) of the GDPR- Principles relating to processing of personal data Principle of purpose limitation

[91] Large-scale processing activities appear in Article 35 (https://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:32016R0679&from=ES#d1e3596-1-19) y 37 (https://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:32016R0679&from=ES#d1e3792-1-1) , and in Recitals 80, 91 and 97. There can be great scale processing activities that do not use Big Data as well as uses of Big Data that do not entail great scale processing activities.

[92] Already a few years ago, the AEPD published a Best Practices Code in data protection for Big Data projects (https://www.aepd.es/sites/default/files/2019-09/guia-codigo-de-buenas-practicas-proyectos-de-big-data.pdf)

- With regard to the duration of the data processing activity
- And with regard to the geographical scope of the processing of *Big Data.*

Notwithstanding any other considerations that may arise out of the specific context of certain data processing operations, as may be the frequency with which certain processing operations may be performed out of data obtained initially.

Thus, by tacking the development of a solution based on *Big Data*, the Administration acting as controller, upon prior performance of an impact assessment or EIPD, must take into consideration a series of facts so as to minimise the risks that the processing may present for the rights and freedoms of individuals, and adopt a series of precautions and guarantees in the design of the different operations that are part of the processing.

- **Phase of data acquisition:** the data processed will need to be minimised through a prior selection of the data that need to be collected and the level of detail with which such data are processed needs to be minimised through the anonymisation or the pseudonymisation of the sources of origin, the masking of the data or the encryption of the information.
- **Analysis and validation phase:** as in the operation above, the detail of the data must be minimised to the extent possible through anonymisation techniques and encryption techniques.
- **Dissociation, anonymisation, or pseudonymisation phase of the information:** preferably, the persons in charge of this activity will not be the ones involved in the information exploitation phase, a recommendation that becomes an obligation for the data controller when such data are health data, as stated in additional provision 17 of the LOPDGDD notwithstanding any other obligation stated in such additional provision, more precisely, the guarantee of the traceability of the information within the frame of the guarantees envisaged in the GDPR. It needs to be highlighted that the very dissociation, anonymisation, or pseudonymisation process is in itself a processing of personal data. Therefore, the guarantees envisaged under the applicable data protection regulations need to be applied[93].
- **Storage phase:** the confidentiality of the data needs to be guaranteed as well as the fact that these are not accessed by unauthorised third parties, and use, to that end, encryption techniques and authentication and access control mechanisms. It is likewise important, for the purpose of avoiding inferences arising out of an unauthorised comparison of the several information sources, to resort to data distribution strategies that avoid making connections between the data.
- **Exploitation phase:** when a use is going to be made of data in order to extract value and present the information arising thereout, the anonymisation of such data needs to be guaranteed through the use of the different techniques available and, as the case may be, the legal guarantees aimed at avoiding reidentification[94], if a prior use thereof has been made and the data in this phase still allow for an identification of data subjects to be made.

The process of aggregation to obtain knowledge entails combining data, often from different information sources, with the privacy risks for data subjects this entails.

[93] Opinion 05/2014 on anonymisation techniques

[94] It is advisable to read publications by the AEPD such as: Orientations and guarantees in Personal data anonymisation procedure, la Introduction to the hash as a personal data pseudonymisation technique o The K-anonymity as a privacy measure.

1. **Re-identification of individuals or singularisation**, augmenting the probability for this to occur the greater the volume of data processed is, even in datasets that may not contain primary or explicit identifiers in principle. [95].
2. **Traceability** of the different registers of a single data subject or group of data subjects, be it in the same dataset of through a connection of sources of data that are heterogeneous and independent, through a correlation analysis.
3. **Inference**, out of personal data classified as quasi-identifiers[96], of personal information much more critical whose processing had not been foreseen.

These risks for privacy must be assessed upon the very ideation of the processing that uses *Big Data* techniques and its analytical tools of information exploitation, and include, by design, the strategies needed to mitigate them and that will have been identified as a result of the impact assessment for data protection that will have been performed.

When information is enriched of the same person with data coming from different sources of information, we may probably discover new connections or subtleties of the personality that would not have appeared on their own. That is, we deduce and derive certain information on the individual. Therefore, regardless of the specific purpose of the processing, it is necessary to analyse the risk that undesired inferences may occur, especially regarding special data categories.

In addition, it is even possible that, at the time to compare several data sources that are allegedly anonymous, due to the aggregation of data, the identity be revealed of certain individuals: out of general traits, the number of individuals in the intersection of all subjects decreases until specific individuals are identified[97]. The risks of reidentification need to be measured, assessed and managed and the necessary measures in order to reduce the probability that such reidentification materialise need to be adopted. Furthermore, the possible reactive measures need to be envisaged to mitigate the possible damage that may arise for a natural person in the event that the reidentification finally occurs. If the data pertain to special data categories, such as health data, we face an even greater risk due to the greater impact that this entails for the rights and freedoms of individuals. Therefore, the guarantees to be adopted need to be higher, a question that needs to be taken into account at the time to process data pertaining to minors or especially vulnerable individuals.

If this analysis is important for the internal use of data by the Public Administrations themselves, it is even more important and must likewise be performed before the anonymised transfer of data by the Public Administrations to third parties. In case the risk of reidentification is high, such processing would fall under the data protection regulation and would need to be lawful.

For the case of the risks associated to the reidentification of data subjects and the inference of the information the following approaches can be of use:

- **Data minimisation:** the processing of data needs to be limited to the maximum extent and reduced to what is necessary to fulfil the purpose of the processing, both from the point of view of the volume of the population (number of entries) and of the data analysed (attributes processed).

[95] The primary or explicit identifiers are those than identify an individual unequivocally, such as their tax number, their social security number or their mobile phone.

[96] Quasi-identifiers are attributes that do no identify the person directly (date of birth, postal code, gender, profession,...) but allow reidentification if combined or compared to other datasets that share these same quasi-identifiers.

[97]The closest simile we can think of is the game for children called 'Guess Who?' (Guess Who?) on Wikipedia. https://en.wikipedia.org/wiki/Guess_Who%3F

- **Maximisation of the level of aggregation:** to the extent possible, we must avoid reidentifying the individuals or inferring information on them within the *dataset.* In order to do this, it is necessary to minimise the detail of the information processed.
- **Abstraction of the information:** the personal data need to be protected and their relation needs to be concealed.
- **Data distribution:** to the extent possible, data should be distributed and processed in backgrounds that are, when not physically separated, at least logistically separated, all this for the purpose of hindering the inference of information through a comparison of the data. Another approach, which is similar to the techniques used in AI, is the use of federated exploitation models[98].

In order to implement these strategies, several techniques may be used. Inter alia, anonymisation techniques[99] (data erasure[100], generalisation[101], perturbation[102], permutation) in case of minimisation and aggregation techniques, and information encryption for data abstraction and distribution techniques[103.] Masking techniques are likewise useful[104].

These strategies aimed at maintaining the anonymisation and the dissociation of data need to be complemented with other measures that seek to guarantee transparency, control by the users of their won data through suitable mechanisms o as to exert their rights and application of the proactive responsibility principle by the data controller through the monitoring, auditing and traceability of the decisions adopted and the actions accomplished.

The risks added to the solutions for which the data processing evolves into a decision-making need not be overlooked[105]. In this case, the same recommendations as those established in the chapter on AI in this document need to be followed.

Last, and even if this is a “classic” of the risks in the use of technologies, the system must be secure and be designed with regard to data protection as a requirement by design and by default. Neither hiring cloud services easily , nor the variety of the tools to perform massive computing or the fact that they may be free or the access to open data can serve as an excuse to propose projects where security and protection are not taken care of regarding the information from one end to the other and throughout the life cycle of the processing in operations where clients, providers, employees and the users themselves are involves.

[98] These models follow an approximation to “take the processing to the data instead of taking the data to the processing”. Thus, problems such as privacy, ownership and localisation of the data can be addressed.

[99] In 2016, the AEPD published the Guidelines and guarantees in personal data anonymisation procedures. https://www.aepd.es/sites/default/files/2019-09/guia-orientaciones-procedimientos-anonimizacion.pdf. We can also find Opinion 05/2014 on anonymisation techniques by the Article 29 Working Party, which specifies these techniques together with other anonymisation techniques, their robustness and guarantees, and typical errors arising as a result of their applicationhttps://www.aepd.es/sites/default/files/2019-12/wp216-es.pdf

[100] Through the erasure, the value of an attribute is completely eliminated (usually, explicit identifiers) or substituted with a pattern (for example, “*”)

[101] Generalisation is a technique that is usually performed on quasi-identifiers and consisting in substituting the values of an attribute for a more abstract and more general attribute within the taxonomy of the attribute (for example, substituting the age with a range of ages or substituting the full postal code that identifies the municipality with the two first numbers that only identify the province)

[102] Perturbation consists in substituting the original values of the data (such as through the addition of noise) in such a way that the connection that may exist between the original registers is suppressed but the statistical properties of the original data are preserved.

[103] Several privacy strategies oriented towards data protection are described in the Guide to Privacy by Design, published by the AEPD.

[104] Data masking is the process through which certain elements of the data in a data warehouse are modified through a change in their information but preserving a similar structure, in such a way that the sensitive information is protected: https://www.powerdata.es/enmascaramientode-datos

[105] Article 22 of the GDPR establishes a general prohibition to adopt automated individual decisions based on sensitive personal data (Racial or ethnic origin, political opinions, religious or philosophical beliefs, trade union membership, genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or a natural person's sex life and/or sexual orientation.) unless the explicit consent by the data subject is obtained, or for public interest purposes, provided that the adequate measures have been adopted to guarantee the rights and freedoms and the legitimate interest of the data subject.

VI. ARTIFICIAL INTELLIGENCE

A. WHAT IS ARTIFICIAL INTELLIGENCE?

The term Artificial Intelligence or AI is associated to machines or information systems that, as part of the implementation of their tasks, are capable of learning from their own experiences and of solving more or less complex problems in different situations, in such a way that they seem to "think" or show a certain intelligence[106].

This intelligence, in a process such as that of adjusting the focus of a camera[107], may apparently lack importance. However, it can also be very sophisticated, such as in the case of a medical diagnosis expert system. In the last years, Artificial Intelligence has left the lab rooms and has become included within multiple daily-life systems, such as Internet browsers, automated translation, smart watches, smart fridges, smartphone apps, and, of course, the smart systems of the Public Administrations.

Under the section of Artificial Intelligence many techniques can be included [108]. We could speak about natural language recognition, automated learning, pattern recognition, etc. Often, the smart system is capable of very complex tasks, such as playing chess or planning logistic routes, but it is seldom used in the context of small decisions that could be easy and repetitive for a human being, thus allowing for a substitution of humans with machines. This is the case for text or telephone *chatbots*[109], to set appointments or keep a closed dialogue with the user, as well as spam filters or the recognition of persons in photographs, and even the recognition of expressions or moods.

Another element that has enhanced the introduction of artificial intelligence in many processes in the way in which applications are developed. If the traditional development of programs is based on defining a process precisely through the specification of a series of elemental instructions, machine learning-based systems[110] use a different technique. This development starts form more imprecise instructions on how to do the work [111], but enough so as to start operating, and the system is "trained" with test data, learns the relations between inputs and outputs, and corrects its results when it misses, obtaining a system that know how to work thanks to its "prior experience". This way to develop systems is more similar to the way in which we learn certain skills, it gets along well with certain tasks that are difficult to define and to specify.

The AI algorithms are designed to recognise patterns, thus learning to take automated decisions. The use of these algorithms for decision-making does not always allow to find the

[106] The High-Level Expert Group on Artificial Intelligence (AI – HLEG) created by the European Commission to develop the European Strategy on Artificial Intelligence, in their document named "A DEFINITION OF AI: MAIN CAPABILITIES AND DISCIPLINES" suggest applying the term Artificial Intelligence to ""systems that display intelligent behaviour by analysing their environment and taking actions – with some degree of autonomy – to achieve specific goals)"

[107] Canon. What is AI-focus https://www.canon.com.au/explore/glossary/ai-focus

[108] The guide "Compliance with the GDPR for processing activities involving Artificial Intelligence. An introduction", in the section "Introduction to the AI frame and data protection", develops, inter alia, a brief description of the AI techniques, the life cycle of an AI solution and the possible processing activities in each of the phases. https://www.aepd.es/sites/default/files/2020-02/adecuacion-rgpd-ia.pdf

[109] Chatbots or conversational boxes. https://en.wikipedia.org/wiki/Chatbot

[110] Machine Learning: is the ability of a smart system to adapt and generate new decision rules out of a series of predefined rules for the purpose of improving its rate of correct decisions.

Deep Learning: a more detailed learning system than Machine Learning that includes several input and output layers that allow to learn the general relation with one another in subsequent steps. Thus, the margin of error is reduced and the precision is augmented of the conclusions, with a lesser level of human orientation being requested.

[111] In Artificial Intelligence backgrounds and aspects related to it such as Machine Learning, the processing of the natural language or the construction of expert systems, logic (or declarative) programming languages are used that describe, through algorithms, the necessary computing logic to solve a problem without describing a control flow of any kind. This paradigm is based on the formula "algorithms = logic + control" (the so-called Kowalski's informal equation), which means that an algorithm created specifying knowledge through axioms (logic) and the problem is solved through a mechanism of inference that acts on the problem (control). https://www.genbeta.com/desarrollo/lenguaje-prolog-ejemplo-paradigma-programacion-logica

trace or explanation to the decisions adopted. This type of scenario is called “black box”[112]. This lack of transparency affects several aspects, such as the way in which decisions are adopted by the algorithm, the influence of each element of information, the consistency between the different inferences, the precision or accuracy of the inferences or the biases introduces, inter alia. It is for this reason that a great interest exists in developing methodologies that allow for an audit to be made on the algorithms that infer new information and adopt decisions directly and indirectly on individuals[113].

The artificial intelligence also raises concerns among users, researchers, specialists, authorities and the industry regarding legal compliance, the guarantee of the data subjects’ rights and the legal security of all intervening parties. The guarantee of the rights and freedoms of data subjects whose data are processed by this type of solutions requires a suitable data governance that covers quality and integrity of the data used, their relevant in the light of the domain where the AI systems are to be deployed, their access protocols, the quality of the algorithms used and the ability to process data in a way in which privacy is protected.

Furthermore, given that the models of the AI systems may be based on great amounts of data to learn and to adopt decisions, when personal data are being processed it is important to understand the way in which these data affect the behaviour of the systems and to guarantee their quality and their integrity. Also, it is necessary that they do not contain biases, inaccuracies, errors or mistakes that are socially constructed and lead the system to extract incorrect generalisations and to the possibility that it may adopt unfair decisions favouring some groups over others. But, mainly, it is essential that a legitimacy exist for the processing of the data. This need of real data may lead the data controller to use data collected for a different purpose and inconsistent to train their system, thus creating a conflict with the lawfulness and the principles of processing

Often, artificial intelligence systems process personal data and even special personal data such as health, gender, race, beliefs, or political opinions. Such processing may extend further from the input data, including inferred data that are obtained out of such information. Thus, the restrictions established by Article 9 of the GDPR and the LOPDGDD need to be kept in mind together with the conditions to set such limitation.

The moment an AI has autonomy in decision-making, it must be taken into account that the GDPR, in Recital 71, as well as in Article 22, limits and establishes rights in the sense that data subjects whose data are being processed be not subject to exclusively automated decisions[114] that originate legal effects or significantly affect the data subject, including automated profiling[115,116]

Another key obligation that needs to be taken into account is the obligation to inform data subjects, which needs to be adapted to each phase of the life cycle of the AI solution where the processing is taking place. In this regard, apart from the information that needs to be provided pursuant to the regulation, cases where the data subject is being subject to

[112] Inability to understand and/or to reproduce the decision adopted by the AI.

[113] In February 2020, the European Commission edited a document called On Artificial Intelligence - A European approach to excellence and trust, (https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en). The Council of Europe, on its part, has expressed its concerns in the document Artificial Intelligence and Data Protection: Challenges and Possible Remedies (https://rm.coe.int/report-on-artificial-intelligence-artificial-intelligence-and-data-pro/16808e6012). The Ministry of Science and Innovation is actively working on an Artificial Intelligence National Strategy. http://www.ciencia.gob.es/stfls/MICINN/Ciencia/Ficheros/Estrategia_Inteligencia_Artificial_IDI.pdf

[114] Article 22 of the GDPR- Automated individual decision-making, including profiling

[115] The Article 29 Working Party has analysed the implications of this right in the Guidelines on automated individual decisions and profiling for the purposes of Regulation 2016/679https://www.aepd.es/sites/default/files/2019-12/wp251rev01-es.pdf

[116] For it to be considered that a human involvement exists, the supervision of the decision needs to be made by an authorised person with the competence to modify the decision, and a significant action needs to be performed that is not just symbolic.

automated decisions or the cases of profiling referred to in Article 22 of the GDPR are of special importance, as in these cases, the data subject must have available the *“significant information on the logic applied” and “the importance and consequences envisaged”.*

As a good practice, and other than the requirements arising out of data protection, a qualified human supervision is advisable in any AI-based processing and, in general, in processing activities where automated decisions are adopted. At the time to design these systems, human supervision is an option that needs to be taken into account so that it is part of the procedures and mechanisms associated to the processing, in such a way that it allows the possibility for a human operator to ignore the algorithm at any given time. Thus, a control element would be introduced that focuses on providing an answer to the possible risk that the decision adopted by the system be incorrect or limiting of the rights and liberties of data subjects.

In addition, it is necessary to establish the procedure to be followed vis-à-vis those situations where this way of action needs to be adopted, for which it is advisable to document any request of human intervention or questioning of the automated decision received by the data subjects, in such a way that it is possible to detect situations out of an analysis on such automated decision and that it is possible to detect situations where this intervention is needed because the model may not be working as envisaged.

We can also find systems that are sending personal data without applying any guarantee on such data to third countries that do not meet the necessary guarantees for the purpose of executing part of the functionalities of a smart system in their servers.

In order to help organisations including AI components in their processing activities comply with the GDPR, the AEPD has published the document RGPD compliance of processings that embed Artificial Intelligence. An introduction detailing aspects of responsibility, legitimacy, exercise of rights, application of the proactive responsibility, etc.

B. THE AI IN THE PUBLIC ADMINISTRATIONS

It is beyond any doubt that the artificial intelligence is a reality that has come to stay, and that it will be increasingly present in our future. The volume and the variety of the data available together with the increase in the process capacity are the breeding grounds for the development of solutions in different areas of activity sectors, inter alia, the public sector[117], for the purpose of automatising the processes and offer more efficient and improved services.

The European Commission, in compliance with its AI strategy, has developed a coordinated plan with State members to encourage the development and the use of AI in Europe. This plan envisages the creation of a group of experts of artificial intelligence (AI HLEG in English, *Artificial Intelligence High-Level Expert Group)* composed of representatives of the world of academics, the civil society and the industry. Their goal is to support the implementation of the European Strategy on Artificial Intelligence taking into account several backgrounds that include data protection, and to work on the definition of a reliable artificial intelligence. In order to reach this reliability, the Commission believes that AI must meet seven key requirements: human action and supervision, technical robustness and security, privacy management and data management, transparency, diversity, non-discrimination and equity, social welfare and environmental welfare, and accountability.

At state level, the Spanish Government is actively working on the creation of an Artificial Intelligence National Strategy whose aim is to line up national policies that seek to encourage

[117] https://analiticapublica.es/ia-y-administraciones-publicas?utm_source=inapsocial&utm_medium=pagado&utm_term=nuevo&utm_content=ia&utm_campaign=transformaciondigital

the development and the use of AI in Spain. The tools for such purposes are the increase on the investments, the reinforcement of the excellence of AI technologies and applications, and the strengthening of the cooperation between the public and the private sector. The goal is for a significant impact to occur in society and in the Spanish economy.

The Spanish R+D+I Strategy on Artificial Intelligence[118] identifies several areas for the development of AI in the public sectors: inter alia:

- Interaction with the citizen, for example, through the use of *chatbots*[119] [120]based on the processing of natural language as the first interface between the citizens and the Public Administration.
- Health, both for the processing and for the management.
- Security in terms of vigilance, mobility and traffic, or, for example, specific subjects such as urbanism inspections
- Corruption prevention.

Even if the use of the word "intelligent" is often associated to a system that is a synonym of the words advanced, careful or customised, thus creating an attractive tag for a specific system, there are parallel cases of failed uses of the AI when the application of AI has been based on expectations that are commercially attractive and have given rise to developments and deployments that lacked a defined purpose or were acquired without the necessary analysis of proportionality and necessity[121]

Any AI component that is developed will not be isolated and will be integrated into a specific processing together with other components, and, thus, personal data may be encountered in the different phases of the life cycle of the solution: training, validation, deployment, exploitation and withdrawal. This is why Public Administrations using this type of solutions need to verify that a minimum set of conditions are met in order to guarantee the compliance of the processing made, the first of such conditions being the existence of a legal ground that may differ for each of the phases of the life cycle of the solution in which the processing of personal data takes places and that may be different.

It must be taken into account that the legal bases, such as the legitimate interest, are banned for the Public Administrations. This does not mean that, if a third party has developed an AI system using such legal basis for the processing built on an AI, this cannot be used by the Public Administration. What will occur is that Public Administrations will not be able to use such legal basis to justify any subsequent processing. The selection of the legal bases is tightly related to the purpose of the processing and the goal sought in each of the phases of the life cycle of a system and, in the case of Public Administrations, with their assigned competencies.

[118] Ministry of Science, Innovation and Universities – Spanish R+D+I Strategy on Artificial Intelligence (2019) - http://www.ciencia.gob.es/stfls/MICINN/Ciencia/Ficheros/Estrategia_Inteligencia_Artificial_IDI.pdf

[119] As part of this type of software or computer programs of artificial intelligence we may find the chatbots, which are bots (BOT: a computer program set to perform repetitive tasks with the use of artificial intelligence) specialised and created to hold conversations and offer preconceived answers.

[121] There are likewise disappointments associated to the 'AI solutionism' (https://towardsdatascience.com/risks-of-ai-solutionism-dangers-of-machine-learning-and-artificial-intelligence-in-politics-and-government-728b7577a243)

C. Risks to rights and freedom associated to the use of Artificial Intelligence in Public Administration

Some of the specific risks of AI are related to the type of development used. For example, if massive use of data is applied to train the systems, as happens with machine learning, the same risks described for big-data based processing must be considered.

Systems often integrate artificial as third-party engines and components [122] embedded in the processing performed by the controller and interact with them. Those embedded components may be as simple as a library or source code or as complex to complete systems devoted to process data in the providers' machines. That is, it is entirely possible to have an information system in a public administration which integrates an artificial intelligence system which is being executed in real time in a server located in Asia, from which it is fed data and yields the corresponding results. These "on the cloud" procedures, or simply, the recourse to third party components and to a chain of data processors, is a source of risks that need be addressed and managed.

Since a processing may be associated to an ever-evolving AI and date change with time, there is a risk that its scope may impact in the accuracy of the inference performed or cause a bias in the results.

Besides, whenever IA models include personal data, those data are under a risk of exposure to third parties as a consequence of an attack of any kind. For this purpose, several techniques may be used, similar to those used for Big Data, and regarding both privacy as security by design, in order to guarantee personal data protection.

If an IA-based processing used decisions solely based in the automated processing of collected data, public entities are advised to analyse any risks arising from this decision-making process and implement measures for analysis and management, such as:

- Establishing committees for ethics and data protection[123] which shall be responsible for assessing potential damages and benefits that a certain processing may entail for society at large, and, most especially, for data subjects.
- To establish periodic quality guarantee controls on its system in order to ensure that people are treated in a fair and non-discriminatory way.
- To carry out audits in order to ensure that the components used in automatic decision-making systems are operating as desired[124].
- In order to avoid entirely delegating the decision on a self-sufficient system and assuming its conclusions without having anybody revising them, and which may

[122] The most popular examples of these AI engines are IBM Watson, Tensor Flow de Google, Amazon Lex or Microsoft Azure Machine Learning. A good comparison of engines can be found at: https://www.agicent.com/blog/best-ai-engines/

[123] The Government of Spain R&D+I Strategy states, in the development of priority 1, the involvement of the Spanish Committee or Ethical Research in the use and implementation of AI. This is further developed in priority 6, where it is highlighted that the Committee or Ethical Research must lead any analysis and assessment activities of ethical aspects on the use and implementation of AI in the activities of the R&D+I plans.

(https://www.ciencia.gob.es/stfls/MICINN/Ciencia/Ficheros/Estrategia_Inteligencia_Artificial_IDI.pdf)

Some initiatives in this sense have already been implemented, such as the Committee for an Ethical Artificial Intelligence created by the Professional Insurance of Attorneys for the financial and insurance sectors (https://www.mutualidadabogacia.com/sala_de_prensa/cronica-comite-etica-presentacion/)

[124] The covid-19 pandemics has obliged to forgo on-site exams British academic authorities were forced to use an algorithm to rate the GCE exams which enable students to access university. Said algorithm, besides considering consistent factors, such as the student academic history, has included other, more controversial, factors, such as the academic history of their school and of other student in their class; this algorithm was heavily criticized when it was proved to be particularly unfair with students from ethnic minorities coming from the less-favoured, poorest environment, whose marks have been weighted down by comparison with other students from more well-off areas. https://www.xataka.com/robotica-e-ia/cuando-nota-no-te-pone-profesor-sino-algoritmo-caos-estudiantes-reino-unido

have a large impact on people's rights and freedoms, it is advisable to introduce the further guarantee of a subjective approach that makes a true connection between data and results.

- Implementing mechanisms that allow the data subject to express their point of view and to challenge the decision, if appropriate, including information on the terms assigned to review and a contact person for consultation.
- Even when an human control obligation[125] is deemed to apply, for those procedures that involve automated decision making, a risk that needs to be assess is that there might be no qualified human operator who cannot ignore the results of the decision taken at a given time. For this reason, it is fundamental to prepare procedures designed for those situations identified as problematic based on incident or case analysis, in which data processing has not acted as expected and which a human supervisor must de facto be used to verify the suitability of the decision taken.

Another risk to be considered is that AI systems may be manipulated to generate erroneous inferences about certain individuals or groups or individuals. These manipulations may have their origin at factory setting and be included as a backdoor that allows subsequent manipulation or attacks by exploiting potential vulnerabilities of the system. Therefore, the need to perform audits for the purpose of detecting such vulnerabilities must be assessed and, if deemed appropriate, their scope and periodicity must be determined.

Another aspect to be considered is that during system operation and "authority" effect may arise among the operators responsible for interacting with the solution forgo any critical review of the decisions or inferences generated by the AI and therefore a really qualified human control does not actually take place, or, even, that automated decisions are revised but recorded as erroneous, creating doubt or biases in future human decision-making processes. The need for human supervision must be understood as a method for continuous improvement of AI processes, which need to involve the entire staff in the relevant processing.

Chatbots, or intelligent dialogue features used in communications channels of Public Administration, may not be accessible for persons with disabilities. In such cases, the risk of not having a natural and rational method, that is, by means of human interfaces, of contacting public entities, must be assessed, studied and solved in order to prevent any risk of discriminative processing resulting from the use of AI.

Chatbots also presents other risks, such as the risk of collecting and disclosing personal information to third parties when they have a continuous learning procedure based on information fed by users. In this same case, if the algorithm is allowed to evolve without supervision, it risks to change its reasoning model towards erroneous inferences. Chatbots may present security vulnerabilities and be uses by hackers to access personal information [126]. They may also collect more information than necessary, include biases affecting data quality or carry out inferences on special categories. Those risks are more persistent when chatbots are used to provide assistance with regard to mental health hazards or to crime or abuse victims and children.

[125] Pursuant article 22.3 of the GDPR, when a contract is executed between the data subject and the data controller, or when a processing is based on the explicitly given consent of the data subject, the data controller shall be responsible to implementing the necessary measures to protect the rights and liberties and the legitimate interest of the data subject and, at the very minimum, the right to seek and obtain human intervention by the controller, to express their perspective and to challenge the decision resulting from the processing.

[126] Its vulnerabilities may be used to implement social engineering attacks: https://venturebeat.com/2017/05/29/what-happens-when-hackers-attack-chatbots/

As for the aforementioned audit [127], considering that assessing the actual behaviour of these systems constitutes a challenge, an analysis of requirement or pre-operation testing do not always suffice - a continuous audit needs to be performed on the results obtained in order to verify that they are appropriate and that they evolve parallel to social change.

[127] Performing audits as a way to verify regulatory compliance by the data controller is explicitly included in article 28.3 of the GDPR on the obligations of the data processor and in article 39.1.b) in the functions of the Data Protection Officer.

VII. BLOCKCHAIN AND DISTRIBUTED LEDGER TECHNOLOGIES

A. WHAT IS A BLOCKCHAIN

Blockchain [128], Distributed Ledger Technologies (DLT) or Bitcoin are terms that have become pervasive over the last years and have raised a great deal of interest lately. Technologist, among other persons, have been surprised by innovations such as virtual currencies such as Bitcoin[129], which apparently manages stored information with transparency and integrity.

Simply put, blockchain may be defined as a network of participants (peers or nodes) who share a distributed ledger in two things are recorded: who has what (assets) and who negotiates with whom (transactions). Unlike traditionally centralized systems in which databases are controlled by a single central authority, blockchain systems all nodes have a copy of this ledger. This makes it extremely difficult to tamper with the information recorded in this ledger without the entire networks (understood as each and every participant) become are of this attempted tampering. The term "blockchain" reflect the manner in which the corresponding ledgers are organized: a series of blocks in which transactions are grouped and which are chained to one another in chronological order, by means of a cryptographic mechanism called hash which guarantees the integrity and immutability of the information registered in the chain.

In fact, blockchain is a particular instance of a wider concept which are Distributed Ledger Technologies (DLT), which described decentralized, distributed databases managed by several participants, usually peers in the sense that they all have these same authorities. Since there is not a central authority that carries out the functions of verifying and validating information, participant consensus mechanisms are established in order to decide how to make decisions, how to update data and how to store information in a consistent manner.

Blockchain promoters believe in a future in which accounts, contracts or ledges do not need to be placed under the care of an intermediary or a custodian. All transactions would be registered in apparently unchangeable and transparent chains, which may even be automatically updated by means of smart contracts [130] inserted in the chain itself.

There are several types of blockchains, but all of them share the same characteristics in varying degrees: transparency, decentralization, integrity and trust. These characteristics are one of its main advantages: they allow to forgo intermediaries in processing transaction between parties. However, this is also one of its great weaknesses when providing legal protection to the relevant transaction.

In order to advance to this scenario, several different blockchain structures have been promoted; they differ on their degree of visibility or the scope of the operations performed on them (public or private) and by the method by which the each blockchain is managed (permissioned or not-permissioned).

As it is the case with any other technology, an implemented blockchain system may exclusively process information which does not include personal data, for example, a record

[128] The Wikipedia article on Blockchain is a good introduction https://en.wikipedia.org/wiki/Blockchain
[129] Bitcoin: A Peer-to-Peer Electronic Cash System. Satoshi Nakamoto https://bitcoin.org/bitcoin.pdf
[130] https://en.wikipedia.org/wiki/Smart_contract

of transportation of goods[131]. However, as soon as a blockchain network is used to leek personal data, it is obvious that any processing in which this method is included needs to comply with the requirements established by the GDPR, which are, among others, to clearly identify the data controllers (any other potential roles), and the legitimate grounds for processing, to offer a guarantee of confidentiality and availability with regard to data and the possibility for data subject to exercise a series of rights.

Another issue to be considered when deploying this type of networks is how are they going to be affected by data location. GDPR is very clear with regard to transfers of data between countries with different levels of protection[132] and the concept of a distributed hosting networks rather contradicts those guarantees. Information encryption algorithms are becoming increasingly robust, but, at the same time, it is being discovered that the alleged robustness of years back can be broken now using increasingly accessible hardware.

Before starting a blockchain project, the compatibility of this solution with the regulation requirements provided by the General Personal Data Regulation needs to be assessed[133]. More specifically, the following aspects need to be carefully analysed;

- **Responsibility for the processing[134]**. The blockchain, although depending on the type of blockchain implemented is, by definition, a distributed system where it is not easy to identify the data controller(s),
- **Right to be forgotten[135] and right to rectification[136].** One of the principles on which blockchain is that it is impossible to alter the contents without creating an inconsistency, and therefore, the existence of solutions of removal or modification of registered information must be carefully assessed.
- **Storage limitation[137].** One of the principles governing personal data processing are that such data must be "kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed"; therefore, alternative mechanisms need to be implemented to solve the inherent immutability of the network.
- **Security[138].** For some blockchain implementations, there are to significant security aspects which may be compromised. The first aspect compromised by placing the information on a network is data confidentiality. The second, and less obvious, is data availability; although, in principle, information is distributed across many nodes, generally those nodes do not offer an availability guarantee (agreements of level of service) or even a commitment to enduring existence. This is especially so in public blockchain network implementations.
- **International transfers of data[139]**. An aspect that needs to be considered is that using blockchains and any technological solutions may, in many scenarios, cause international transfers of data, due to the very nature of this technology. This is especially so in public blockchain networks. A good way of dealing with this risk is

[131] Regulation (EU) 2018/1807 of the European Parliament and of the Council of 14 November 2018 on a framework for the free flow of non-personal data in the European Union exists and is fairly updated (https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018R1807&from=ES)

[132] Chapter V of the General Data Protection Regulation is entirely devoted to international transfers of data.

[133] Blockchain and Data Protection in the European Union, Michèle Finck, Max Planck Institute for Innovation and Competition; Oxford University, February 2018 https://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID3119584_code1137858.pdf?abstractid=3080322&mirid=1

[134] Article 24 of the GDRP: Responsibility of the controller

[135] Article 17 of the GDPR: Right to erasure

[136] Article 16 of the GDPR: Right to rectification

[137] Article o 5.1.e) of the GDRP: Principles relating to processing of personal data. Storage limitation.

[138] Article 32 of the GDPR: Security of processing

[139] Article 44 of the GDRP: General Principle for Transfers

to apply privacy-from design principles, carefully choose the type of network to be used (considering hybrid or private blockchain networks) and its governance model (for example, off-chain storage of data).

B. Blockchain technologies in Public Administration entities

The practical applicability of blockchain, besides cryptocurrencies and environments without legal security, has found problems for deployment. This is the case for Public Administration entities when they act within the scope of their competencies. Their actions are defined but the existence of a regulating legal frameworks and the acknowledgement of a trusted third party, which is an approach that significantly different from blockchain.

However, a large number of initiatives[140] and concept testing for potential use of blockchain both in corporate and governmental environments have been developed. With regard to the latter, in many places of the work several pilot projects have arisen such as ownership registries, titles and offer registries, traceability of foodstuff products, or identity management or electronic vote.

As stated above with regard to other technologies, when its use involves publishing own contents by the corresponding Public Administration entities, which may include personal data either form private citizens and from public employees, the provisions set forth by the GDPR must be complied with.

On 10 April 2018, UE Member States plus Norway, subscribed a joint declaration in order to create the European Blockchain Partnership (EBP)[141] and agreed to cooperate to establish a European Blockchain Services Infrastructure (EBSI) which would be in charge of supporting the delivery of services to the trans-border public digital maintaining the highest security and privacy standard. Certain governments[142] and public entities are developing blockchain-based projects or strategies, such as:

- **Academic certifications:** Give citizens back control for managing their academic certifications, significantly reducing verification costs an improving trust in authenticity, and thus preventing fraud in official, proprietary and private certifications.
- **Notary Public services.** Blockchain capacities can be used to create shared registries of digital auditing records, to automate compliance verification in time-sensitive processes and to prove data integrity.
- **Self-managed European Identity:** Allows users to create and manage their own identity across border without having to rely on a centralized authority. The goal of this project is to allow citizens to own and manage their own digital identities and to decide who may have access to which of their data.
- **Reliable data exchange.** Using blockchain technologies to allow UE customs and tax authorities to exchange data in a secure manner.

IN 2019 the European Commission deployed a pilot network to test all used cases by means of a user story titled “Eva’s Journey”[143] developed on the European Blockchain Services Infrastructure (EBSI), and new use cases are being developed.

At a national level, different pilot project and concept testing are being carried out in order to analyse the potential use of this technology in the public sector and to identify the specific

[140] EU-Funded Projects in Blockchain Technology https://ec.europa.eu/digital-single-market/en/news/eu-funded-projects-blockchain-technology

[141] European Blockchain Services Infrastructure (EBSI) - https://ec.europa.eu/cefdigital/wiki/display/CEFDIGITAL/EBSI

[142] Malta, Germany, Australia, United Kingdom, Finland, Sweden, have projects deployed on blockchain-based technologies.

[143] A good explicative video can be found at https://www.youtube.com/watch?v=m2uj7fgb2JI

cases in which its use could effectively improve the services provided to private citizens, including:

- **Academic certification** In the context of the project on Academic Certifications developed within the EBSI, the CRUE-TIC, through the BLUE (Blockchain network of Spanish universities) is developing a concept test to record all academic certifications, as well as the competences and skills associated to them, which shall enable students to manage the exchange of such certifications with any stakeholders in the sector.
- **Electronic registry of powers of attorney:** The City Council of Bilbao us developing a blockchain network and an app to manage the electronic registry of powers of attorney. This network and app would be interoperable with all other registries from Public Administration entities, in compliance with article 6 of Act 39/2015, of 1 October, on Common Administrative Procedures for Public Administration. It is based on the existing platform already deployed by the IT Company of the Basque Government, EJIE (Eusko Jaurlaritzaren Informatika Elkartea), which, at the end of 2017, assigned a contract aimed to the development and implementation of a blockchain solution designed for scenarios related to the registration of contractors of the Basque Public Administration entities.
- Bid registry and automated assessment: it the same line, the General Directorate for Procurement, Equity and Organization of the Government of Aragón is developing a distributed registry of bids and a service of automated assessment of such bids, to be used in electronic procurement procedures, which was authorised at the end of 2018 with the goal of increasing transparency and security on public procurement procedures.

Blockchain is but one of the elements of a processing, but does not replace all other organizational and technical elements of such proceeding. As happens in all projects that develop a new service of a new implementation of a service, must be submitted to suitability, proportionality and necessity tests[144]. If, after a previous analysis regarding feasibility, technical and regulatory compliance , the relevant Public Administration entity decides to develop an application on service based on blockchain technology, the desired data processing model must be thoroughly assessed, understanding is context, its scope and its implications for the purposes of identifying and assess any possible threat to the rights and freedoms of natural persons.

As for DLT technologies based projects, Regulation (EU) No 910/2014 of the European Parliament and the Council of 23 July 2014 on electronic identification and trust services for electronic transactions in the internal market and repealing Directive 1999/93/EC, known as the eIDAS Regulation and Royal Decree-Act 14/2019, of 31 October, adopting urgent public security measures in the field of e-administration, public procurement and telecommunications.,

The latter was enacted for the purpose of guaranteeing public security related to the used of any identification and electronic signature systems, and established that all systems included under section 2 of articles 9 and 10 of Act 39/2015, of 1 October, on Common Administrative Procedures for Public Administration, that is, those base on a personal keyword an any other system deemed valid by the Public Administration under the established terms and conditions to have users identified before Public Administration entities, must necessarily comply with the requirement that the technical resources needed to collect, store, process and manage such systems, are located with the territory of the European Union, and within the Spanish territory for data of special categories in the sense

[144] Blockchain Technology Overview. Dylan Yaga, Peter Mell, Nik Roby, Karen Scarfone. In page 42 of this same document there ins a interesting flow chart on the feasibility of blockchain-based projects https://nvlpubs.nist.gov/nistpubs/ir/2018/NIST.IR.8202.pdf#page=53

of Article 9 of the GDPR. Those data, with the exceptions provided by law, cannot be transferred to a third country or international organizations and, in any rate, shall be available for access for the competent judicial and administrative authorities.

Additionally, article 3 of the aforementioned Royal Decree-Act includes a sixth additional provisions to Act 39/2015, of October, which provides that identification systems based on Distributed Ledger Technologies and signature systems based on the latter shall not be admissible in any circumstances in the context of the relationship of private citizens with public entities, and therefore cannot be authorised for such purposes, unless they become the subject of specific national regulations which are in turn compliant with the European Union legal framework. Besides, this sixth additional provision establishes that any identification system based on Distributed Ledger Technologies provided by national laws must take into consideration that the General State Administration shall act as an intermediate authority, play win whatever role is necessary, to guarantee public security.

C. RISKS TO RIGHTS AND FREEDOMS ASSOCIATED TO BLOCKCHAIN TECHNOLOGIES

Before any blockchain-based solution requiring personal data processing is implemented, it is necessary to value, in the framework of an impact assessment regarding data protection, the necessity y and proportionality of using this technology with regard to other possible alternatives and the additional risk introduced due to its own conceptualization and design.

Once blockchain technologies have been selected as the chosen technological solution, another important decision regards the type of network which is to be implemented as processing support. If a processing project compliant with the GDPR is already implemented, the risks entailed by switching to blockchain systems other than a private blockchain network managed and under total control of a Public Administration entity needs to be assessed.

Implementing a blockchain-based system in the services provided requires Public Administration entities to carry out a proportionality assessment, and to balance the main benefits for services provided against the main challenges for data protection mentioned above. It is the responsibility of the public Administration to assess whether the technological solution adopted with the most appropriate solution or, on the contrary, it introduces new risks that cannot be managed. Particularly, it is important to assess which type of blockchain network architecture (public or private, permissioned or non-permissioned) is best adapted to the specific solution, for not all of them entail the same level of risk; equally important is to define the most appropriate governance model. It must be considered that the impossibility of technically complying with any of the obligations of the data controller set forth by the GDPR must not be considered a risk to be addressed but a non-compliance.

When the chosen blockchain is supported by a governing mechanism which is not under the control of the Public Administration, it must be considered that changes may be made and that those changes may entail a hazard to data protection rights and principles. Besides, it must be determined whether there is a risk that the service quality levels regarding node availability or response times are not complied with. In an extreme situation, all the nodes in a blockchain network may stop working at once. For this reason, besides the corresponding legal guarantees, contingency plan must be set up in order to guarantee processing continuity, and which may include, for example, portability options in order to assure that the rights of the data subjects may be protected.

The blockchain security is fundamentally supported by the robustness of its cryptographic mechanisms. All cryptographic systems have an undetermined but limited life cycle. Therefore, obsolescence and the possibility of having the entire system compromised at a certain time need to be managed.

The nature of the distributed system makes more complex to perform data access controls and to keep an access register. This may be aggravated by the fact that the blockchain model chosen allows dynamic input and output of participants.

With regard to this aspect and to the dynamic participant configuration in the blockchain, the chosen model of consensus mechanism must be managed when generating the block chain and the tampering risk must be continuously assessed.

Pursuant to the principle of data protection by design[145], a pseudonymisation strategy is to limit the information fed to the blockchain network to commitments . Should this not be possible, hashes[146] should be used by means of a keyed-hash function or at least, as encrypted hashes, so that the confidentiality of the information recorded in the chain is guaranteed. It is important to avoid disclosure of any identifiable personal data in the block chain; such data should be kept off-chain, in traditional databases or any other information systems maintained by the controller.

It must not be forgotten that block chain is the technological solution, or one of them, that are used to support data processing; and that, in the framework of such data processing, there may be other tools to access and manage data or even to provide user intermediation mechanisms. All those elements have their own vulnerabilities, especially with regard to security; associated risks must be analysed and the necessary measures need to be implemented in order to prevent identity theft for execution any transaction [147] [148] [149] [150]and guaranteeing integral security.

[145] Article 25 of GDPR: Data protection by design and by default

[146] The analysis "Introduction to hashes as a personal data pseudonymisation technique" published by the Spanish Data Protection Agency, explains what a hash function is, its use as a pseudonymisation technique, as well as the hazards it entails and its use limits. https://www.aepd.es/sites/default/files/2020-05/estudio-hash-anonimidad.pdf

[147] https://es.cointelegraph.com/news/cisco-and-ukrainian-cyber-police-uncover-50-mln-bitcoin-phishing-scam

[148] https://criptomonedaseico.com/noticias/los-ciberdelincuentes-generan-bitcoin-utilizando-la-estafa-falsa-de-la-pagina-web-de-la-bbc/

[149] https://www.criptonoticias.com/seguridad-bitcoin/detectan-campana-phishing-robar-bitcoins-basada-dominio-blockchain-info/

[150] A. A. Andryukhin, "Phishing Attacks and Preventions in Blockchain Based Projects," 2019 International Conference on Engineering Technologies and Computer Science (EnT), Moscow, Russia, 2019, pp. 15-19, doi: 10.1109/EnT.2019.00008. https://www.researchgate.net/publication/333072391_Phishing_Attacks_and_Preventions_in_Blockchain_Based_Projects

VIII. SMART CITIES

A. WHAT ARE SMART CITIES

The concept of smart city[151] has been used by the European Commission to group a series of very different technological projects with the common goal of optimizing resource management in cities[152] by means of using technological solutions to manage more efficiently the services typically provided in a city, such as transportation, infrastructures, water, gas and power supplies, waste management or social welfare.

The most characteristic projects for smart cities are based, first of all, in automated data collection by means of different sensors places throughout the city to monitor traffic, movements of persons, air quality, power consumption, etc. Then such data are automatically analysed; this can be made by means of integrated methods and every enriching data with data from other sources. If those data are appropriately combined, they may be used to forecast and dimension which services need to be offered. Finally, there is a conclusions or decision-making stage which may, in some cases, be applied automatically by means of actuators; for example, to regulate road traffic in a big city. This application may involve other technologies described in this document, such as artificial intelligence.

In consequence, although it is possible to speak about smart cities, this term in fact designated the integration of other different technologies, such as artificial intelligence and big data techniques (both also described in this document), and most particularly, the development of IoT (Internet of Things) projects[153] with the purposes of managing a city.

B. SMART CITIES AND PUBLIC ADMINISTRATION

Inasmuch as it is the function of the Public Administration entities to provide citizens with quality services as efficiently and sustainably as possible, smart city technologies offers local governments tools to obtain real-time information with regard to certain services, behaviours and city dynamics, by means of sensor or data sources. Such data sources may be smart pedestrian counters, mobile phone use data, transport card use data, among many others.

Recently, IESE Business School (associated to the University of Navarre) published the sixth issue of the *Cities in Motion Index* (CIMI)[154] corresponding to year 2019. This study assesses data from 176 of the main cities in the world [155] for the purposes of ranking and comparing, among other things, the use of quality-of-life-improving technologies Some of the real-live applications of these technologies in actual cities are:

- **Energy efficiency:** by means of cameras of sensors that detect whether there are any persons in the building and consequently turn on the lights or the A/C system.
- **Efficient management of traffic and mobility:** by installing cameras and sensors at strategic points, so that they can measure the persons and vehicles in the area and prioritize ones or the others in function of their volume and need.
- **Waste disposal by means of smart waste systems:** Sensors are used to detect the level of waste in waste containers and notify a trash truck of when to empty a particular bin; this reduces the noise pollution caused by trash trucks and reducing

[151] Again, Wikipedia is a good source of information in terms of information and references. https://es.wikipedia.org/wiki/Ciudad_inteligente

[152] European Commission's Smart cities/Smart living website https://ec.europa.eu/digital-single-market/en/smart-cities

[153] An introduction to the Internet of Things can be found in Wikipedia https://en.wikipedia.org/wiki/Internet_of_things

[154] The IESE Business School – Cities in Motion can be accessed in the following site: https://media.iese.edu/research/pdfs/ST-0509-E.pdf

[155] Several Spanish cities are included; their respective rankings are given in brackets: Madrid (24), Barcelona (28), Valencia (61), Seville (76), Malaga (80), Palma de Mallorca (88), Zaragoza (101), A Coruña (102), Murcia (105), Bilbao (107).

costs by being able to develop more efficient routes adapted to the actual level of waste in each container.

The last example shows that not all data processed by smart cities programmes are personal data. Besides, compliance with efficiency goals of smart cities does not justify a massive, indiscriminate collection of personal data. The minimization principle[156], collected data must be

adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed..

The purpose of the collection may be either executing public functions, scientific research or statistical purposes. In such cases, it must be assessed whether these processing methods are compatible with the aforementioned purposes, as established in articles 5.1.b and 89 of the GDPR, provided that they are subject to appropriate guarantees. For example, when information us automatically collected in real time for statistical purposes, among other legal, technical and organization measures, the sample size must be assessed, as determined by the number of subjects or elements in a population needed to ensure that the data obtained are representative and effective. The amount of personal data collected, understood both in terms of number of subjects as the dataset collected from a single subject, have to undergo proportionality and necessity tests.

In this sense, for many data processing associated to smart city strategies it does not seem necessary to univocally identify citizens; the same purposes may be served using anonymous aggregated data. In order to forecast when and where traffic peaks occur, or when and where power supply or waste collection need to be optimized, it is not necessary to identify the actual persons that benefit from those services.

Public Administration, in respect of the principle of minimization and considering the necessity and proportionality of the processing, must assess whether they desire to identify private citizens, or whether for the intended purposes aggregated and anonymised data may suffice. In order to forecast when and where traffic peaks occur, or when and where power supply or waste collection need to be optimized, actual identification of the actual persons that benefit from those services must be avoided.

Another aspect to consider is the principle of fairness[157] understood as guarantee that the collected data are to be used for the intended purpose sand goal.. For example, whenever the purposes of a processing are to guarantee that access restrictions to a certain part of a city are implemented, it is important to separate the purposes of access control from the purposes of imposing penalties to those that do not comply with the corresponding restrictions. Therefore, the relevant processing must pursue maximum efficiency in the original purposes without ceasing to be proportional with regard to the degree of intrusion; maximum efficiency alone shall not do.

Managing a series of projects such as those associated to smart cities may require the collaboration of different Public Administration entities With private companies and other agents with whom data are exchanged. This exchange requires interoperable format and data accessibility.

Another question which needs to be solved in projects like this, in which several agents carry out each a part of the treatment, is who makes the relevant decisions and thus acts as the data controller[158]. It is also important to set out who are the joint controllers[159] and

[156] Article 5.1.c of the GDRP: Principles relating to processing of personal data Data minimisation.

[157] Article 5.1.a) of the GPDR: Principles relating to processing of personal data Lawfulness, fairness and transparency.

[158] Article 24 of the GDRP: Responsibility of the controller

[159] Article 26 of the GDPR: Joint controller

processors[160] and how the relationship between those roles is structured. More specifically, it is important to define which data are used for each role and for which purposes.

In all cases in which personal data are collected, and regardless of whether such data are subsequently anonymised, data subjects must be appropriately informed[161] in the framework of the service in which their data are collected, which of their personal data are collected and for which purpose and how can they exert their rights over them. The website of the relevant Public Administration entity may be a good place to publish this information and, very importantly, to keep it updated.

A core element of a smart city projects are city dwellers. Providing these city dwellers with appropriate information regarding collection of use of their data, beyond strict compliance with articles 13 and 14 of the GPDR and application of privacy by design as provided by Whereas 39, allows individual persons to be aware of any risks, regulations, protections and rights associated to the processing of their personal data, as well as of the channels to exercise their rights with regard to such processing; this in turn increases their level of trust an commitment with the relevant project. The website of the relevant Public Administration entity is also a good place to raise awareness about specific data-sharing projects involving other public or private entities, and therefore clarify and explain the added value that is intended to be obtained by this project. Open data and data sharing request by third parties must be compliant with the principle of transparency that make processing lawful.

C. RISKS TO RIGHTS AND FREEDOMS ASSOCIATED TO SMART CITIES

When preparing a contract to be executed by controllers, joint controllers and processors, it is important to consider that Smart City projects are long-time projects, and that their consequences may evolve greatly depending on changes occurring at social or technological levels. Those contracts must consider aspects such as maintenance of devices and systems, their ownership and the responsibility to update them, as well as support to the data controller to manage all risk to right and freedoms (probably including a DPIA).

Public Administration entities must be aware that, the larger the amount and the higher the frequency of information collected, the higher the inherent data protection risks that individual citizens are subject to for this reason. Processing of this kind are, by their own nature, high-risk[162] in most cases due to the large amount of data that are amassed by Public Administration entities For this very reason, the Guide for Local Administrations[163] highlights that, before deploying a smart city project, the following assessments need to be made:

- Previous assessment of the project with regard to the volume of information that needs to be processed, the number and types of sources from which this information is to be obtained, data collection frequency and the time during which this information is to be kept.
- Data enriching assessment, both planned within the processing and as a risk.
- An impact assessment of the terms established by Article 35 of the GDPR assessing even the need, depending on project characteristic, of escalating a previous consultation to the Spanish Data Protection Agency.

[160] Article 28 of the GDPR: Data processor

[161] Articles 13 and 14 of the GDPR - Information to be provided where personal data are collected from the data subject and information to be provided where personal data have not been obtained from the data subject, respectively.

[162] Article 35.3.c) of GDPR: Data protection impact assessment (a systematic monitoring of a publicly accessible area on a large scale)

[163] Guide for Local Administrations(AEPD): https://www.aepd.es/media/guias/guia-proteccion-datos-administracion-local.pdf

Governance of a smart city system must integrate compliance for problem solving as set forth by a relevant data protection policy. Besides the specific risks inherent for a smart city, such project must assess and manage the risks associated to the involved technologies.

Even when the smart city project is based on data anonymised at origin, their extension, frequency, combination and enrichment may cause a person to be (re)identified by them. We all have our unique habits, which reveal our job, our partner(s), our health status and even our political ideas and religious beliefs. The risk of re-identification must be assessed and mitigated by measures such as the application of differential privacy techniques, use of aggregation strategies to prevent correlations, using local and distributed process into reduce to amount of data stored centrally under the same data controller, etc.[164]

Massive deployment of sensors and actuators increases the probability of security breaches in any of the three relevant domains: confidentiality, availability and integrity. Therefore, a very delicate aspect to be considered is the security against possible breaches and malicious[165]. This is what happened on a massive scale in 2007 in Tallinn[166]. A security risks analysis carried out from a data protection approach must provide the maximum possible guarantees to prevent non-authorised access that allow to monitor individual persons or a massive leak of persona data, which is one of the biggest risks had by smart cities.

Although security can never be complete, preventive plans of continuous audit can be implemented, such as ethical hacking, risk evolution assessment (for example, by deploying new services and technologies) and measures that minimise the impact that a breach, however improbable, may have. These measures involve applying data criteria, separation of data in time and in different systems and categories, or hiding data by means of encryption, early pseudonymisation or blocking techniques. In some cases, it may be interesting to allow data subjects to exert control on the automatic collection of their data, implementing autonomous and transparent mechanisms that consider citizen participation in such a way that they can monitor that relevant guarantees are complied with.

There are two critical moments in which rights and freedoms are at the most peril: deployment and removal. Deployment means that the systems is not yet adjusted, especially with regard to data protection risks; therefore, it is convenient to make limited deployments and to seek assistance by control authorities. Removal or replacement of the entire system has a system abandoned or almost abandoned, which may cause an additional vulnerability.

[164] As stated before, reference should be made to the following Spanish Data Protection Agency publications: Guidelines and guarantees in personal protection anonymisation procedures, Introduction to hashes as a personal data pseudonymisation technique or K-anonymity as a privacy measure..

[165] Smart Cities Cyber Security Worries. 2018 IOActive, Inc. Graph in https://ioactive.com/wp-content/uploads/2018/10/IOActive-SmartCities-cybersecurity-worries.pdf

[166] This event caused a reaction at the NATO: https://elpais.com/diario/2007/05/18/internacional/1179439204_850215.html

IX. CONCLUSIONS

Digital transformation of the Public Administration is a reality arising from an intensive use of new technologies, some of them state-of-the-art innovations, which pursues both improving internal operations and providing a better service to users. Possibilities offered by new technologies in sectors such as healthcare, education or investigation are endless, and there is little doubt that new technologies are a key element to improve efficacy and efficiency of public services.

Notwithstanding the fact that all services provided by Public Administration entities have a public service purpose, processing of personal data within these organization involves an inherent risk arising from the sheer volume of data subjects, of the extension of collected data, of the impossibility, often, of opposing to processing and of the power imbalance between the Public Administration and the individual citizens.

Inherent risks may involve, for data subjects, in case of breaches of the rule of law, abusive situation by public decision-makers, massive or selective filtering of personal data, security breaches, changes in law even in third countries, emergencies, events gone out of control, etc.

The fundamental right to data protection arises from the need to prevent the tragic events of the 20th century which were possible due to an extensive [167], intrusive [168] and, in some cases, automated to a certain degree [169] processing of personal data, even despite the personal effort of public servants which fought against a system which clearly overpowered them[170].

In this sense, it is important to remember that data protection regulations are a tool that, beyond protecting the rights and freedoms of citizens, allows emerging and innovative technologies to be used in Public Administration entities in a manner which is both sustainable and compliant with the social responsibility of the Public Administration.

The data protection standards in Europe and Spain are modern and is well aligned with good engineering practices and current project management. This is, basically, a proactive approach in the sense that it addresses problems before they occur, highlights the responsibility of the different agents, and enforces continuous risks management aimed at prevention, early detection and planned response to risk. In this sense, it can be said that compliance and common sense go hand in hand.

Technology is changing the world, and it is important to design services and applications that guarantee that this change is for the better, at least with regard to personal rights and freedoms. Introducing technology in the processes traditionally carried out by Public Administration entities, or having such Public Administration entities implement new entities based on technology, is increasingly simple, but should not be increasingly risky, intrusive or uncontrolled[171]. Quite the contrary, creating new services in an ever-evolving society, both at social and technological level, involves a higher risk exposure that needs to be assessed and managed on an individual basis and forces data controllers to better analysed the consequences of those considering that things do not always turn out as expected.

[167] Dutch Civil Registries in 1940 included information on the religion of each citizen when they were seized by the German invading forces: https://en.wikipedia.org/wiki/History_of_the_Jews_in_the_Netherlands#The_Holocaust https://en.wikipedia.org/wiki/1943_bombing_of_the_Amsterdam_civil_registry_office

[168] Data collected by the German Democratic Republic: https://elpais.com/diario/1991/10/07/internacional/686790013_850215.html

[169] IBM and the Holocaust: The Strategic Alliance Between Nazi Germany and America's Most Powerful Corporation. Editor: Dialog Press 2012 ISBN-13: 978-0914153276

[170] Brief biography of René Carmille: https://hipertextual.com/2018/12/rene-carmille-hombre-que-hackeo-nazis-ii-guerra-mundial

[171] Although old, the project management triangle or iron triangle is still fully valid. The sentence "good, fast, cheap: choose two" is widespread. https://en.wikipedia.org/wiki/Project_management_triangle .

Public decision makers must also consider an additional factor: that technologies are not accessible to all[172]. This does not only apply to communication networks, computers or tablets. Many people do not understand technology,, for the reason of being educated in other times, of their own limitations or of not being interested enough or having enough time to educate themselves. Even if all legal requirements to inform them about what is to be done with their data, this may not suffice: it is important to guarantee that the information is provided in such a way that it can be understood by them.

None can protect an asset if they do not know they have, or if they ignore that it is valuable and that its loss or misuse can cause damages. One of the responsibilities of Public Administration, but also of each and every data controller, is to educate their users in why they need to be careful with their data, how are those data to be used by third parties and how they will benefit from letting the Public Administration use their data[173].

This document assesses a series of more or less state-of-the-art technologies, and highlights certain risks that Public Administration entities, acting as data controllers, must consider if they adopt said technologies as a support for the data processing they carry out. It is important to highlight that this analysis is not complete or exhaustive in any way, neither as an exhaustive list of risks or as a complete range of technologies. On the contrary, its purpose is to be used as an example of the analyses that data controllers need to perform whenever they decide to include a new solution in their digital transformation process within their area of competences: suitability, necessity and proportionality assessments, identification of risks introduced by them, impact assessment with regard to data protection and interference with compliance of other obligations stated by data protection standards.

A proactive management of data protection related risks on the part of the Public Administration, together with adherence to appropriate criteria of transparency, proportionality, minimization and restriction of processing function as key factors to guarantee compliance of services provided and generate trust among citizens.

[172] The digital divide is the unequal distribution in access, use or impact of Information and Communication Technologies (ICT) across social categories. These categories may be based in gender, territorial, geopolitical, cultural or other factors. https://en.wikipedia.org/wiki/Digital_divide

[173] The May 2018 Sociological Survey carried out by the Spanish Centre of Sociological Investigation asked interviewees on personal data protection and the potential use of personal information by third parties (question 9). 37,3% reported that this was a "source of great concern" to them and 38,8% described it as a "source of concern". (http://datos.cis.es/pdf/Es3213mar_A.pdf)

X. ANNEXES

A. USEFUL REFERENCES AND RESOURCES

- Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)

  https://www.boe.es/doue/2016/119/L00001-00088.pdf

- Organic Act 3/2018, of 5 December, on Protection of Personal Data and Guarantee of Digital Rights

  https://www.boe.es/boe/dias/2018/12/06/pdfs/BOE-A-2018-16673.pdf

- Guide to Privacy by Design

  Spanish Data Protection Agency, October 2019

  https://www.aepd.es/sites/default/files/2019-11/guia-privacidad-desde-diseno.pdf

- Guidelines for Data Protection by Design

  Spanish Data Protection Agency, October 2020

  https://www.aepd.es/media/guias/guia-proteccion-datos-por-defecto.pdf

- Guidelines for Managing Security Breaches

  Spanish Data Protection Agency, June 2018

  https://www.aepd.es/sites/default/files/2019-09/guia-brechas-seguridad.pdf

- Practical guidelines for Risk Analysis in Personal Data Processing

  Spanish Data Protection Agency, February 2018

  https://www.aepd.es/sites/default/files/2019-09/guia-analisis-de-riesgos-rgpd.pdf

- Practical Guidelines for Impact Assessment in Protecting Data Subject to GDPR Regulations

  Spanish Data Protection Agency, October 2018

  https://www.aepd.es/sites/default/files/2019-09/guia-evaluaciones-de-impacto-rgpd.pdf

- Report Template for Data Protection Impact Assessment (DPIA) for Public Administration

  Spanish Data Protection Agency,

  https://www.aepd.es/sites/default/files/2020-03/modelo-informe-EIPD-AAPP.rtf

- List of processing types requiring a DPIA (art. 35.4)

  Spanish Data Protection Agency, September 2019

  https://www.aepd.es/sites/default/files/2019-09/listas-dpia-es-35-4.pdf

- Non-exhaustive list of data processing that do not require an impact assessment regarding data protection (art. 35.5)

  Spanish Data Protection Agency, August 2019

  https://www.aepd.es/sites/default/files/2019-09/ListasDPIA-35.5l.pdf

- Guidelines on Cookie Use

Spanish Data Protection Agency, July 2020

https://www.aepd.es/media/guias/guia-cookies.pdf

- Device fingerprinting analysis

Spanish Data Protection Agency, February 2019

https://www.aepd.es/sites/default/files/2019-09/estudio-fingerprinting-huella-digital.pdf

- Report on privacy policies in Internet

Spanish Data Protection Agency, October 2018

https://www.aepd.es/sites/default/files/2019-12/informe-politicas-de-privacidad-adaptacion-RGPD.pdf

- Decalogue to adjust Internet privacy policy to the GPDR provisions

Spanish Data Protection Agency, October 2018

https://www.aepd.es/sites/default/files/2019-09/decalogo-politicas-de-privacidad-adaptacion-RGPD.pdf

- Guide to comply with the duty to inform.

Spanish Data Protection Agency, Catalan Data Protection Authority, Basque Data Protection Agency, January 2017

https://www.aepd.es/sites/default/files/2019-09/guia-modelo-clausula-informativa.pdf

- Adjustment of processing incorporating Artificial Intelligence to the to the GPDR provisions. An introduction

Spanish Data Protection Agency, February 2020

https://www.aepd.es/sites/default/files/2020-02/adecuacion-rgpd-ia.pdf

- Blockchain and the General Data Protection Regulation.

European Parliamentary Research Service, EPRS, July 2019

https://www.europarl.europa.eu/RegData/etudes/STUD/2019/634445/EPRS_STU(2019)634445_EN.pdf

- Guidelines for Data Protection and Local Administration

Spanish Data Protection Agency, April 2020

https://www.aepd.es/sites/default/files/2019-09/guia-proteccion-datos-administracion-local.pdf

- Guide for customers that subscribe to cloud computing services

Spanish Data Protection Agency, September 2018

https://www.aepd.es/sites/default/files/2019-09/guia-cloud-clientes.pdf

- Guidelines for cloud computing service providers

Spanish Data Protection Agency, September 2018

https://www.aepd.es/sites/default/files/2019-09/guia-cloud-prestadores.pdf

- Code of good practices for data protection in big data projects

Spanish Data Protection Agency, May 2017
https://www.aepd.es/sites/default/files/2019-09/guia-codigo-de-buenas-practicas-proyectos-de-big-data.pdf

- Guidelines and guarantees in personal data anonymisation procedures

Spanish Data Protection Agency, 2016

https://www.aepd.es/sites/default/files/2019-09/guia-orientaciones-procedimientos-anonimizacion.pdf

- Opinion 05/2014 on Anonymisation Techniques

  Article 29 Data Protection Working Party, April 2014

  https://www.aepd.es/sites/default/files/2019-12/wp216-es.pdff

- Introduction to hashes as a personal data pseudonymisation technique

  Spanish Data Protection Agency, November 2019

  https://www.aepd.es/sites/default/files/2019-11/estudio-hash-anonimidad.pdf

- K-anonymity as a privacy measure.

  Spanish Data Protection Agency, June 2019

  https://www.aepd.es/sites/default/files/2019-09/nota-tecnica-kanonimidad.pdf

You can find more resources and references to technology-related materials in the Innovation and Technology Area in the Spanish Data Protection Agency website.

B. CONTACT WITH DATA PROTECTION AUTHORITIES

The Spanish Data Protection Agency is the national independent administrative authority in charge of supervising the implementation of the Organic Act on Personal Data and Guarantee of Digital Rights and the GPDR. Besides, the Autonomous Regions of Andalusia, Catalonia and the Basque Country have their own regional data protection authorities, which are authorised to exert the functions and powers reserved to control authorities by the GPDR, in the cases provided by article 57.1 of the Organic Act on Personal Data and Guarantee of Digital Rights, that is:

a) Data processing the controller of which is a Public Administration entity in the corresponding Autonomous Region, a Local Entity of the corresponding territory or any other organization directly or indirectly providing any services.

b) Data processing carried out by natural or legal persons for the purposes of exerting a public function within the capacities of the corresponding Regional or Local Administration.

c) Data processing provided for in the corresponding Statute of Autonomy.

Pursuant to the above, in general terms Public Administration entities must notify any security breaches to the Spanish Data Protection Agency by means of their website. However, in Andalusia, Catalonia and the Basque Country, when there is a breach of security as those described above, the competent control authority shall be the relevant regional authority, that is:

- In Andalusia: The Andalusia Transparency and Data Protection Council which may be accessed at its electronic portal
- In Catalonia: The Catalan Data Protection Authority (https://apdcat.gencat.cat/es/inici/) at its electronic portal.
- In the Basque Country: the Basque Agency for Data Protection at e-mail address avpd@avpd.eus