# GDPR: enough or too much for IoT?

Ordieres-Meré J.(*) and Jeunemaitre A.(**)

(*) Full professor at Industrial Management Department of the Universidad Politécnica de Madrid and visiting professor at Institut Interdisciplinaire de l'Innovation i3 Centre de Recherche en Gestion. École Polytechnique de París

(**) Director of Research – Management - Institut Interdisciplinaire de l'Innovation i3 Centre de Recherche en Gestion- CNRS (UMR9217)

## ABSTRACT

The paper discusses the extent to which the currently enforced General Data Protection Regulation (GDPR) provides for enough flexibility in fostering the Internet of Things (IoT) data sharing, particularly the way it would benefit companies and society at large.

Following on the introduction the academic literature on the subject is reviewed. It anchors the issue in the current debate about data security, privacy, and value creation. It does so paying special attention to arguments about consistency, reliability, completeness and noise disturbances of the available data.

In investigating the issue the adopted methodology is inductive. It revolves around present-day models on IoT data transfer identifying their limitation. Alternatives are discussed whereby IoT data may be retrieved or erased in great numbers from usage. Exploring the alternatives the paper comments on existing figures from the GDPR publications confronting IoT business value creation processes with regulation.

Finally, the paper addresses specific situations out of the very basic ones to check when such usage could raise privacy issues.

The contribution intends to highlight both the regulatory challenges ahead and how much technologies may provide with solutions together with beneficial opportunities of IoT dissemination 'knowledge' for citizens.

## INTRODUCTION

This is the age where data is seen as gold or 'the new oil' [1], and personal related data is always valuable for companies trying to understand or integrate human behavior in their models, products or systems. However, personal data are 'personal', which means that in the EU zone a set of regulations started to be adopted. This is the case for the General Data Protection Regulation (GDPR)[2], which is about data privacy and the protection of personal data. Interestingly, some new dimensions have been opened in the regulation, like the 'personal data portability'[3].

The GDPR has very specific rules, and these are especially required when a specific type of personal data processing is involved, because it could lead to a high risk from the data subject rights and freedoms perspective. This is very relevant especially when new technologies are considered, as GDPR fines in case of data breaches or non-compliance can be very high (up to €10 million, or 2% of the worldwide annual revenue of the prior financial year, whichever is higher)[2].

The coverage of GDPR is extensive and, as far as it states in its article 7 that "the principle of confidentiality should apply to current and future means of communication", this includes the Internet of Things. Moreover, the article 5 poses specific safeguards in machine-to-machine communications.

While not all Internet of Things applications are about personal data, like for instance the more Industrial Internet of Things (IIoT) applications, it happens in many other use cases. It is clear that as the market share for IoT solution is growing worldwide (including Europe), this is a relevant aspect that requires detailed management. Just to detail some examples, a digital transformation of healthcare is happening, for instance, there is a rapid growth of wearables and connected medical devices that enable remote health monitoring [4]. Healthcare data are extremely sensitive data, then under the scope of the GDPR. Then, there is smart metering regulation whereby personal data on household energy consumption patterns is involved. Another example from the smart home applications can involve safe monitoring of pattern behavior for elder people [5]. In all these cases and many others it is needless to say that data can be personal and it could benefit people but, different risks are present.

Different treats face this context and, from the other perspective it is a nurtured field to be used as source for large benefits for both, companies and individuals. Some aspects were foreseen by the regulation and some others are still to be better identified. However, it is needed to discuss in which ways the IoT environment need to progress as well as how the evolution of the regulation can affect to it, by creating opportunities and challenges.

Therefore, the remaining structure of the paper looks for a Section 2 devoted to formalize the state of the art from both the regulatiorial side but also the technological side. The next Section 3 will dive deep into the properties of the IoT proposed models. The Section 4 will discuss to what end alternative solutions can play a role in such environment. Then, the Section 5 will discuss about integrative solutions that could challenge the regulation and potential effects. Finally, some conclusions will be elaborated.

# STATE OF THE ART

Cambridge Analytica and Facebook story has demonstrated the very real business implications of data usage failure or misuse. From devastating reputational damage to the estimated US$ 60bn lost by Facebook in the immediate wake of the data scandal, it became an excellent example of the implications of the mismanagement of personal data and compromised data privacy [6].

One of the key aspects in the GDPR nowadays is the consistency, starting from the liability dimension. It is requiring to have the appointment of a Data Protection Officer, to be the responsible figure for all the actions related to the personal data handling in all the processes of the company. Then, it becomes compulsory to develop the Data Protection Impact Assessment (DPIA) [2], where Data assets should have been mapped, business context diagrams created and data flows defined. Essential processes should be in place to respond to the fundamental requirements of GDPR, like consent, the right to be forgotten, but firms should also be aware of the potential requirements for the future regulations, like *ePrivacy*.

Although it can be argued that IoT is not concerned by such high level principles, the truth is that as far as such devices are providing data in a yet continuous or discrete (event driven) way, and merging real and virtual worlds, it is partially true. The issue is to what end such data flows are related to personal identities. Therefore, the DPIA makes fully sense, as some processes are related to such data and they are responsible for attaching them to such identities.

Because of the strict deployment of the EU GDPR, some of the organizations scare of failures in properly addressing all the actions needed and they decided to shut down services to EU citizens [7]. However, although it seems to be a defensive approach maybe risking some businesses, it is not fully consistent, except when data processing is located out of the EU zone or products or services are going to be provided out of such area, and more importantly, it is not because of the GDPR but because of the Privacy Shield.

However, the interest of the paper is not just to discuss the GDPR as a whole, but the implications for the IoT and related businesses. The IoT need to be understood not just as a technological innovation making it possible to convert many sensors in active emitters of what they sense and, therefore, being aware of such data streams. Actually, there are proposals [8] not only to get sensor data distributed but also computation capabilities, control and storage services over the cloud-of-intelligent things, like cars [9], domestic appliances, etc. The major motivation for such developments is to provide resources with low latency answer (below few tens of milliseconds), as for car-to-car or car-to-roadside interactions.

The emergence of IoT aims to improve quality of life by connecting smart devices, applications and technologies, empowered by the availability of the high volume of smart sensors. The IoT development started recently to demonstrate applications added value in different areas [10], like Energy, mobile applications, safety, eHealth, Self-quantifying, among others. A significant part of such added value is because they are able to bring new opportunities of business when it works in connection with the "big data"(BD) techniques, which are also growing these days [11-12]. The recent explosive publications of big data studies have well documented the rise of big data and its ongoing prevalence. Different types of BD have emerged and have greatly enriched spatial information sciences and related fields in terms of breadth and granularity. Studies that were difficult to conduct in the past time due to data availability can now be carried out. However, BD brings lots of "big errors" in data quality and data usage, which cannot be used as a substitute for sound research design and solid theories [13]. This is the way the IoT contributes the most to the BD, as it brings additional information to the related aspects, making it possible to reduce the amount of noise in the datasets. Indeed, IoT can also provide multidimensional perspective to the interesting aspects, which can extend the value of the analysis.

However, despite its growth potential IoT still faces many challenges, mainly related to addressing, routing, standardization, restricted physical capabilities (energy, processing, and memory), security, privacy and openness. Therefore, with a bunch of IoT consumer devices, which sometimes are hackable, there is an initial risk to be mitigated, which is to make such hacking activities much less evident, there are strong efforts for developing new security frameworks like the Z-Wave Alliance [14].

Whether usage of consumer IoT devices and data in the consumer-oriented business or when having IoT use cases in an Industrial Internet context whereby personal data is leveraged (e.g. healthcare) with other types of connected devices, there are significant risks under consideration, like unwanted access to the devices themselves, connectivity issues (from the short-range ones such as Zigbee or those used in smart home apps to the many wireless ones in a long-range context, such as LPWA technologies), but also those related to the managing platforms, as clouds where data is supposedly to be stored in a safety way. All those steps are concerned by the GDPR and they must be under control, in such a way that data breaches need to be reported if personal data are involved and under specific conditions (personal data breach notification).

Although these are clear flaws potentially damaging several of the current business, in this paper the interest is to discuss to what end adopted solutions are robust enough and in which way the adopted strategies in the GDPR for enabling the operation of such devices can be enough or they will become an intrinsic limitation.

# IoT VALUE PROPOSALS AND BIG DATA ANALYTICS

The IoT promises significant potential economic benefits. However, current IoT applications are in their infancy and the full potential of possible business opportunities is yet to be discovered [15]. Based on the nature of the application, the IoT devices will result in big or real-time data streams. Applying analytics over such data streams to discover new information, predict future insights, and make control decisions is a crucial process that makes IoT a worthy paradigm for businesses and a quality-of-life improving technology. However, realization of such business opportunities will require consistent business models helping their setup and development. IoT data can be generated quite rapidly, the volume of data can be huge and the types of data can be various.

Most of the existing value proposals for the IoT deployment includes one platform to manage the physical devices in an integrated way, as well as collecting the data from them by placing it into a private cloud to start its processing. Actually, benefiting from RFID and sensor network technology, common physical objects can be connected, and are able to be monitored and managed by a single system [16]. The whole framework can be found in Figure 1.

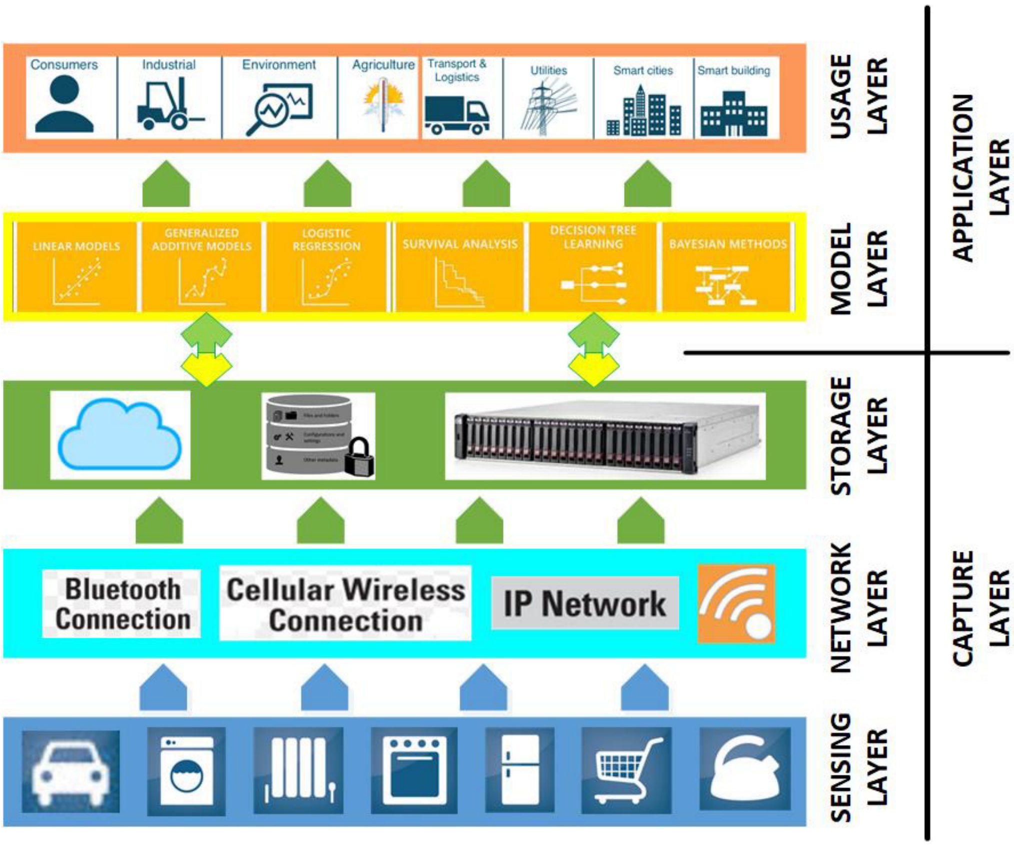


Figure 1.- Architecture framework from IoT upto usage. Several interactions can happen between layers.

The usability of the data coming from the IoT devices enables the BD processing, and each BD sample typically involves a large number of attributes, requiring high-performance servers with a large-scale memory and a powerful computing unit, to cluster and processing big samples. Currently, big data sets are generated and collected from many areas such as social networks, scientific computing and IoT. Although it is obvious, this way of doing also reveals the intrinsic risk derived of such collecting data procedures, which is the mash-up of different sources of data, aiming to create an extensive domain of knowledge, able to improve the foreseen capabilities.

This context benefits the most from the technological evolution of the IoT devices in terms of cost and sensing capabilities, as depicted in Figure 2 and Figure 3.

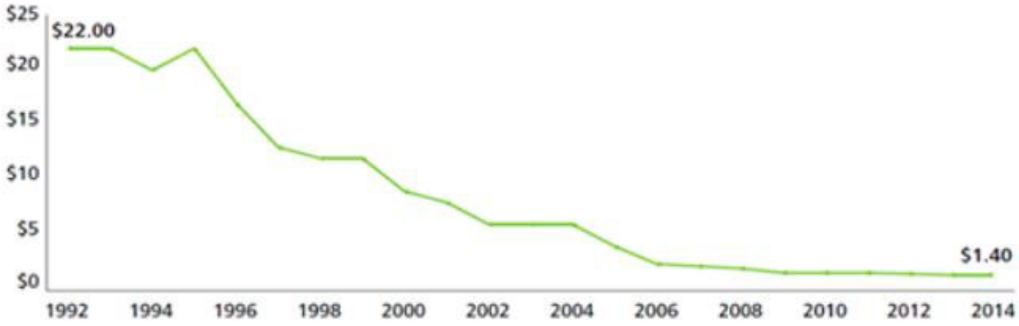


Figure 2.- Cost evolution of prices for sensors. Accelerometer case. Source: https://www.bbvaopenmind.com/ iot-implantacion-y-retos/

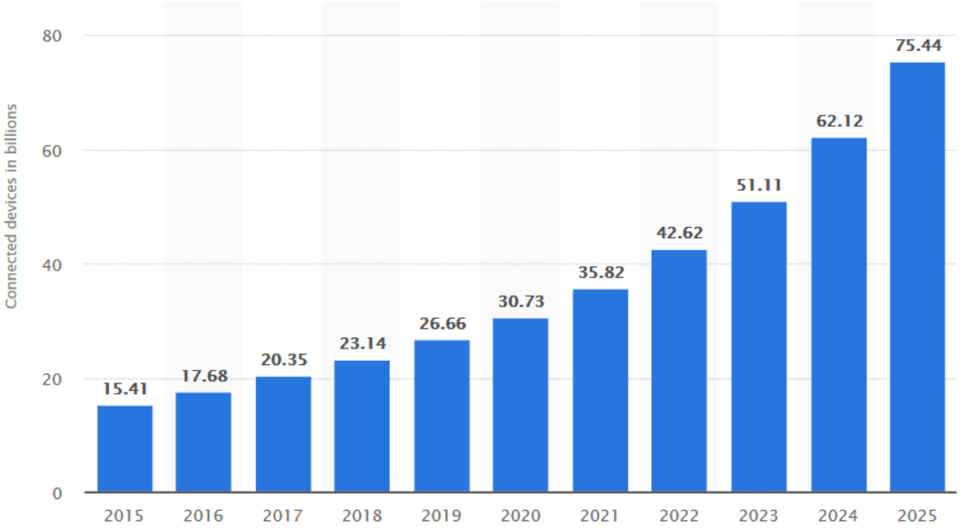
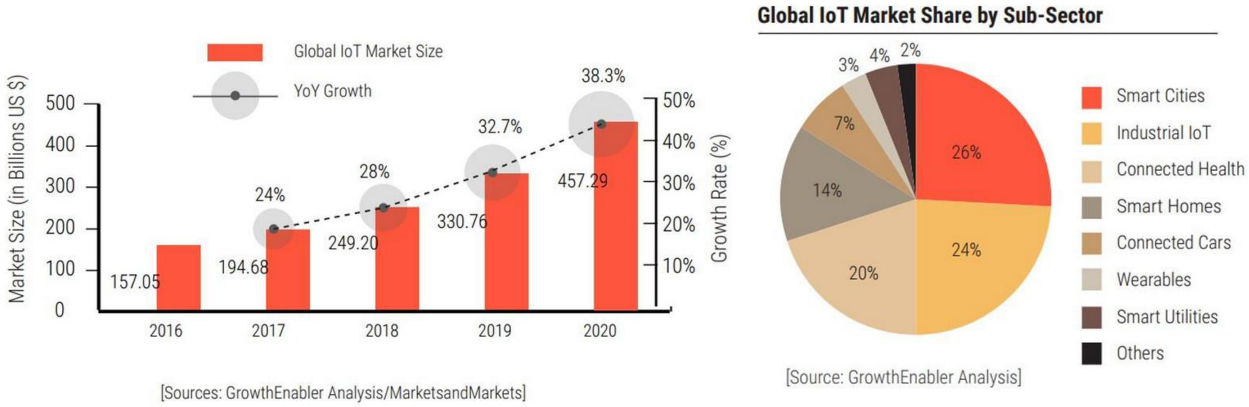


Figure 3.- Growth of connected IoT devices through time. Source: https://www.statista.com/statistics/471264/ iot-number-of-connected-devices-worldwide/

As there is aneed to increase the sources of data and the cost of the devices continue to reduce, it is foreseen an unprecedented growth rate of connected devices. This effect represents yet another point of exploitation in the online world and, despite the promise of the IoT to offer consumers, and society in general, huge benefits including connecting cars, appliances and mobile devices to deliver convenience and personalized services, they are still concerned about their privacy risks. Only 22% of U.S. internet users and 18% of British internet users felt that the benefits of smart devices outweighed any privacy concerns. Privacy concerns could be a barrier to growth, impacting business opportunities that stem from the IoT landscape such as delivering personalized, value-added services.

A Gartner study [17] recently predicted that there will be 26 billion connected devices by the year 2020, but latest research shows consumers want  more information and controls before purchasing or using a smart device.

If we analyze the different segments of market representing the growth, as depicted in Figure 4, it becomes clear how related those markets are to people behavior (Connected health is ranked the third and Smart Homes is the fourth, for instance).



[Sources: GrowthEnabler Analysis/MarketsandMarkets]

[Source: GrowthEnabler Analysis]

Figure 4.- Market of IoT devices through time per sector of usage. Source: https://www.forbes.com/sites/louiscolumbus/2017/12/10/2017-roundup-of-internet-of-things-forecasts/#f2e1a441480e

When asked, 85% of U.S. internet users compared with 83% of British users would want to understand more about data being collected before using smart devices. Additionally, 88% of U.S. internet users and 87% of British internet users would want to control the data being collected through smart devices. Such research shows that consumer privacy concerns (see Figure 5), could hinder the growth of the IoT market as the required trustworthiness is far from being approached.

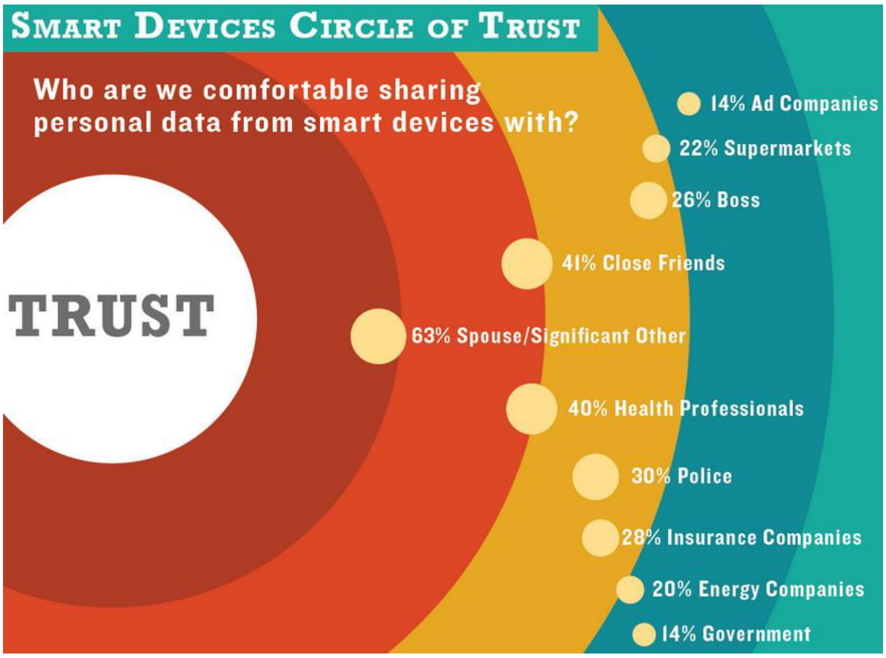


Figure 5.- Circles of Trust for Smart Devices for US citizens. Source: Privacy & The Internet of Things: The Importance of Transparency in Accounting for What We Can't See. https://www.trustarc.com/blog/tag/iot-summit/

Under this perspective, ***GDPR can be seen as effective tool to add significant trust to this market, as the users can now understand their rights and privacy have been increased***. This is another aspect to be deeply discussed.

# GDPR AND PERSONAL DATA COLLECTION

A major aspect of the GDPR are the so-called ***legal grounds for lawfully processing personal data***. One of them is consent. In several IoT applications where consent is used, it might even need to be explicit consent. Others include the specific regulations regarding the processing of personal data regarding children (ample of IoT toys nowadays), the right to be forgotten and the right of access to personal data. However, the latter is is part of the post-consent stages, typically looking at it from the data security, enterprise information management (storage, policies, governance, etc.)

However, as explicit consent will certainly not always be the path to follow, it is necessary to consider the second way established on the GDPR, which is the anonymisation/pseudo-anonymisation. Those strategies are highly recommended by the GDPR regulation, because of if it can be proven that the true identity of the individual cannot be derived from "anonymised data", then this data is exempt from other methods ensuring the strict confidentiality of the actual data.

A similar situation happen with the pseudo-anonymisation, which enhances privacy by replacing most identifying fields within a data record by one or more pseudonyms. Formally, the GDPR defines pseudonymization in Article 3, as "the processing of personal data in such a way that the data can no longer be attributed to a specific data subject without the use of additional information." Then, the "additional information" must be "kept separately and subject to technical and organisational measures to ensure non-attribution to an identified or identifiable person."

There are still after the GDPR some problems with personal data, as the following:

- the data consistency and the use of different sources of information and or different time periods or geographical positions.
- the data storage, although properly scrambled, masked or blurred, the related persons probably are not aware such pieces of information do exist related to them.
- the problems related to EU citizens when requesting products or services outside of the EU-zone.

The harder one is the second one, mainly due to the absence of accountability as you never know who is gathering information about you, then you can not ask for access, erase of modification.

Just to make a clear example how weak the the conditions enforced by the GDPR are, let us describe a potential situation which is, under the good practices enforced by the regulation and still it fails in protecting the personal data. Imagine you ask the phone company operating the phone network to get access to the set of pseudo-anonymized records of the phones registered in different dates in different cities, because we are interested in studying the customer profiles. Imagine you get the ID of the connection tower, the time of entrance for services, the time of exit , and the pseudo-anonymized IMEI of the customer phone.  Formally, it is possible to accept that such records are aligned with the requirements of the GDPR.

However, if selected carefully the cities and the interesting time windows, it will be possible to identify the pseudo-anonymized set of IDs for a football team, or, when refined adequately, you will be able to identify the pseudo-anonymized ID phone of Mr C. Ronaldo. From here, you can start to analyze his specific movements. Then, in the end you have found a personal identity based on the pseudo-anonymized information by combining extra information from other sources.

Although this is just an example, it reveals to what end it becomes difficult to keep the GDPR principles when higher dimensions of data/knowledge become involved. Then, the information discovered can become a keystone to follow up exploiting additional knowledge.

Based on the previous analysis and potential problems, it is less than clear that by hiding the collected data additional protection is given. Actually, it seems it happens in the opposite way. Therefore, to better protect the data the best way is to exhibit them, enabling increasing the awareness.

# DISRUPTIVE PROPOSAL FOR IoT STORAGE AND EXPLOITATION

Based on the previous examples and issues, the future needs to conceive the ***Privacy by Design*** (PbD)[18], the concept of building, embedding security and privacy controls into connected products and infrastructure themselves. Indeed, accountability needs to be dramatically enforced to address the requirements of the regulation.

There are two types of Accountability: Accountability regarding users, but also Accountability within the organization. Another layer happen when accountability is having an independent third party review and verify that your actual privacy practices are consistent and comply with stated practices. Although a third party seal is a good outward indicator for company commitment to privacy requirements, it is even better adopt privacy by design with public exhibition of the collected data duly encrypted.

Apparently this is kind of a paradox but such framework is technologically possible, therefore, it is just matter of redesign the adopted business model [19-20], highlighting this new perspective. In the following paragraphs a potential technical solution will be depicted.

As far as the dominant paradigm is not to encrypt the data streams but keep them hidden from the public either anonymized or pseudo-anonymized, it becomes legitime to think what the situation could be if the data streams become public and encrypted. Therefore, each of us will be aware of the information describing the set of our gadgets and the trust will increase immediately. This because we will be aware about the amount of data related to us, and it will be our responsibility to ask the access to them or to apply for their removal, etc.

Fortunately, there is a technology based on the Digital Ledger Technology (DLT) enabling such kind of approach, we will analyze deeply. It was named as IoTAⓇ it is an open-source DLT designed to be scalable with zero-cost transaction fees and data integrity to power the future of the IoT, machines and identity solutions. Scalability and feeless transactions are mandatory for any identity-related solution and real-world mass adoption. These factors have been the major bottleneck for blockchain-based DLT. The IOTA Tangle, based on Directed Acyclic Graph (DAG) and Masked Authenticated Messaging (MAM), has emerged as a powerful solution to address the real-world challenges of today.

# CONCLUSIONS

The analysis of the implications for the GDPR for the IoT business shows that trying to protect personal data aspects for operations carried out inside the EU-zone is much more than just trying to avoid individual references and masking such references. Although the primary reference for this subject is the controversial between anonymization and pseudo-anonymization and how restrictive they become, we have shown that, based on triangulation with other existing knowledge distributed both in time and space, will add further layers making such data related to specific individuals, which is perfectly supported by the current formulation of the GDPR standard.

Further than how the pseudo-anonymization can be treated, there are aspects not well managed as the rights to modify some contents or to apply for data destruction, if necessary. They are nowadays based on the owner's request, however it becomes hard to do when you are not fully aware of their existence, as produced by decenes of devices and appliances. To deal with this dimension a specific proposal is made in this paper, which is to dramatically change the approach and bring the data stream visible with content encrypted, in such a way the users can browse datasets related to themselves and give appropriate access rights to them, when adequate. Based on the IoTAⓇtangle technology it becomes possible to implement such kind of solutions, including channel subscriptions when interesting the access to continuous data streams.

By means of such kind of changes in the underlying model to collect data, not only the benefit will be the higher awareness of the people to whom the data are referred to, but it can allow developments for micro monetization of such contents, which can create incentives to accept the IoT deployment. Indeed, because of the proposed change, it is expected to increase the levels for the IpT circle of trust.

Therefore, after the analysis, the conclusion is that the GDPR is not a limitation for the IoT development but an extraordinary opportunity to enable new business models based on the IoT collected data, when it becomes developed under a different schema, where awareness of affected users increase and benefits can be linked to them as well.

# REFERENCES

[1] Datoo, A. Data in the post-GDPR world. *Computer Fraud & Security*, vol 9, 17-18, 2018.

[2] Goddard, M. The EU General Data Protection Regulation (GDPR): European regulation that has a global impact. *International Journal of Market Research*, *59*(6), 703-705. 2017.

[3] De Hert, P., Papakonstantinou, V., Malgieri, G., Beslay, L., & Sanchez, I. The right to data portability in the GDPR: Towards user-centric interoperability of digital services. *Computer Law & Security Review*, *34*(2), 193-203. 2018.

[4] Zheng, X., Vieira Campos, A., Ordieres-Meré, J., Balseiro, J., Labrador Marcos, S., & Aladro, Y. (2017). Continuous monitoring of essential tremor using a portable system based on smartwatch. Frontiers in neurology, 8, 96.

[5] Saralegui, U., Antón, M. Á., & Ordieres-Meré, J. (2017, October). An IoT− based system that aids learning from human behavior: A potential application for the care of the elderly. In MATEC Web of Conferences. EDP SCIENCES, 17 AVE DU HOGGAR PARC D ACTIVITES COUTABOEUF BP 112, F-91944 CEDEX A, FRANCE.

[6] Tuttle, H. (2018). Facebook Scandal Raises Data Privacy Concerns. Risk Management, 65(5), 6-9.

[7] Hern, A., Belam, M. LA Times among US-based news sites blocking EU users due to GDPR. Theguardian (28/5/2018). URL: https://www.theguardian.com/technology/2018/may/25/ gdpr-us-based-news-websites-eu-internet-users-la-times. Last visit on 22/10/2018.

[8] Chiang, M., & Zhang, T. Fog and IoT: An overview of research opportunities. *IEEE Internet of Things Journal*, *3*(6), 854-864. 2016.

[9] He, W., Yan, G., & Da Xu, L. Developing vehicular data cloud services in the IoT environment. *IEEE Transactions on Industrial Informatics*, *10*(2), 1587-1595. 2014.

[10] Lee, I., & Lee, K. The Internet of Things (IoT): Applications, investments, and challenges for enterprises. *Business Horizons*, *58*(4), 431-440. 2015.

[11] ur Rehman, M. H., Chang, V., Batool, A., & Wah, T. Y. Big data reduction framework for value creation in sustainable enterprises. *International Journal of Information Management*, *36*(6), 917-928. 2016.

[12] Cai, H., Xu, B., Jiang, L., & Vasilakos, A. V. IoT-based big data storage systems in cloud

computing: Perspectives and challenges. *IEEE Internet of Things Journal*, *4*(1), 75-87. 2017.

[13] J. Liu, J. Li, W. Li, and J. Wu, “Rethinking big data: A review on the data quality and usage issues,” *ISPRS J. Photogramm. Remote Sens.*, vol. 115, pp. 134–142, 2016.

[14] Stojkoska, B. L. R., & Trivodaliev, K. V. (2017). A review of Internet of Things for smart home: Challenges and solutions. Journal of Cleaner Production, 140, 1454-1464.

[15] Sun, Y., Yan, H., Lu, C., Bie, R., & Thomas, P. (2012). A holistic approach to visualizing business models for the internet of things. *Communications in Mobile Computing*, *1*(1), 4.

[16] Jiang, L., Da Xu, L., Cai, H., Jiang, Z., Bu, F., & Xu, B. (2014). An IoT-oriented data storage framework in cloud computing platform. *IEEE Transactions on Industrial Informatics*, *10*(2), 1443-1451.

[17] Pettey, C., & van der Meulen, R. (2013). Gartner says it's the beginning of a new era: The digital industrial economy. *Orlando, Florida*.

[18] Danezis, G., Domingo-Ferrer, J., Hansen, M., Hoepman, J. H., Metayer, D. L., Tirtea, R., & Schiffner, S. (2015). Privacy and Data Protection by Design-from policy to engineering. arXiv preprint arXiv:1501.03726.

[19] Dijkman, R. M., Sprenkels, B., Peeters, T., & Janssen, A. Business models for the Internet of Things. *International Journal of Information Management*, *35*(6), 672-678. 2015.

[20] Zhang, Y., & Wen, J. The IoT electric business model: Using blockchain technology for the internet of things. *Peer-to-Peer Networking and Applications*, *10*(4), 983-994. 2017.