

GDPR: What's in a Year (and a Half)?

Ana Ferreira ^a

CINTESIS - Centre for Health Technology and Services Research, Faculty of Medicine of Porto, Porto, Portugal

Keywords: General Data Protection Regulation, Literature Review, GDPR Compliant Solutions, Privacy and Security.

Abstract: This paper aims to investigate, with a literature review, how the research community has been tackling the security and privacy requirements mandated by the General Data Protection Legislation (GDPR), over the last year and a half. We assessed what proposed solutions have been implemented since GDPR came into force, if and where they were tested in real settings, with what technologies and what specific GDPR requirements were targeted. No similar review has been found by the authors as works in the literature mostly provide recommendations for GDPR compliance or assess if current solutions are GDPR compliant. Results show that most proposed solutions focus on Consent, PrivacybyDefault/Design and are assessed on IoT and healthcare domains. However, almost none is tested and used in a real setting. Although it may be still early days for this review, it is clear that: a) there is the need for more GDPR compliant novel solutions, tests and evaluations in real settings; b) the obtained knowledge be quickly shared so that proper feedback is given to the legal authorities and business/research organizations; and c) solutions on privacy must integrate socio-technical components that can face, in an all-inclusive way, infrastructures, activities and processes, where GDPR must apply.

1 INTRODUCTION


The GDPR (General Data Protection Regulation) (GDPR, 2019) is in force since May 2018. One year (and a half) later, how is the research community tackling the main privacy and security issues to comply with this legislation?

There has been some progress in the general user's expectation, knowledge and awareness on data protection, which has risen in the past year (Breitbarth, 2019) (Alizadeh, 2019). Slight improvement has also been detected in the design of application's permissions and data privacy (Momen, 2019), however, many research results show that, for instance, mHealth applications are not yet ready to comply with GDPR. Some recent examples are (Benjumea, 2019) and (Muchagata, 2019). This domain is a very important sample, as technology in healthcare collects and integrates some of the most sensitive and private data (e.g., special data category – Art. 9 of GDPR). If those are not well taken care of, we can only deduct that other domains are potentially even less protected.

Besides the lack of appropriate privacy policies, with the clear purpose for data collection (Alizadeh, 2019), obvious terms and conditions and the request for too many personal user data to install and use mobile applications, there are still other issues related to the websites themselves, that need to be properly corrected.

This study (Vlajic, 2018) shows that there is overwhelming evidence of widespread and highly covert user tracking in a range of different children-oriented websites. The majority of the discovered tracking is in direct conflict with GDPR since it is performed without parental consent, and by third-party advertising and tracking companies.

On the whole, many problems pertaining to data privacy, data erasure, data transparency, consent, data pseudo anonymisation and other requirements, still need to be adequately addressed. Businesses and other organizations are trying to keep in the race of GDPR compliance but this has been done very slowly. In fact, it may take, not only, more adequate technologies to better face the requirements but also a change in both businesses and consumers/end-

^a <https://orcid.org/0000-0002-0953-9411>

users' attitudes and culture. Turning the direction of what has been done for many years, does take time.

This paper aims to assess whether current research is tackling the issue of testing and implementing, in real settings, the necessary tools and solutions to comply with GDPR. For this, a literature review on research and development on this topic, for the past year and a half, has been performed. An extraction of what are the main problems to be solved, technologies used and proposed solutions, is obtained, with the main goal to provide the research community with useful insight on what still needs to be done, and if current direction needs to be altered or yet, if other directions need to be added to existing ones.

The next section gives a brief overview of GDPR main requirements, key challenges and other reviews that can give some insight on last year's GDPR enforcement. Section 3 presents the methods used to perform the literature review, while section 4 presents the obtained results from that review. Section 5 discusses results, together with some recommendations on how to proceed, and what directions to follow in the near future, for GDPR compliance. Section 6 concludes the paper.

2 GDPR

2.1 GDPR Requirements

The GDPR - General Data Protection Regulation is a legislation applicable to all EU country members and their citizens. Being a regulation, it applies directly to the legislation of each EU country, without much adaptations, unless with the necessary language, cultural and meaning clarifications. The main GDPR requirements and key challenges from the user/citizen's perspective, are the following:

- **Consent (Arts. 6 & 7):** should be obtained from the data owners, using a clear language with clearly defined purposes, before data processing can begin;
- **Transparency (Art. 12):** privacy policies should be easy to access and understand;
- **Special Data Categories (Art. 9):** biometrics data, race, ethnic origin, politics, religion, genetics, health, sex life and sexual orientation, are data that require processing under exceptional conditions, and consent is mandatorily explicit;

- **Right to (Arts. 15, 16, 18, 21 & 17):** the data owner has the right to access data, oppose processing, update/correct data, limit processing, and the right to be forgotten (erase/delete data concerning him/herself);
- **Portability (Art. 20):** transport data between different organizations in proper adaptable formats;
- **PdD (Art. 25):** privacy by design and default – data privacy implemented and taken care from the beginning of data processing, until it is deleted; moreover, **minimisation** of personal data processing regarding the amount of data as well as the period of time that data are processed, is mandatory;
- **Security (Art. 32):** provide adequate technical and administrative measures for personal data protection (confidentiality, integrity and availability), during the whole processing period; this also applies to physically structured data files;
- **Records of Processing Activities (Art. 30):** monitoring and accountability measures for all personal data activities comprising: actors, actions performed, when and what data were accessed/processed (e.g., access logs);
- **DPIA (Art. 35):** perform data privacy impact assessment, whenever required, especially when processing data from special categories or using novel technologies, with very high risk and impact on personal data privacy.

This summary of GDPR key challenges can help identify, within the literature, the main areas of GDPR that have been focused since this regulation came into force. It is possible to associate what are the main proposed solutions and methods to face them and how these have been applied into the real settings.

2.2 GDPR Reviews

To make sure that the review presented in this work had not yet been performed, search queries with the terms “GDPR review” and “General Data Protection review” were applied to online research databases IEEE Xplorer, SCOPUS and ACM digital library, in November 2019.

After the application of the queries on those databases' search engines, the results were: IEEE

Xplorer = 9 articles; SCOPUS = 71; and ACM = 0. After reviewing titles and abstracts of the obtained lists, only one article from SCOPUS was included to be analysed as a full-text. After this analysis, it was concluded that the review did not focus on the main subject of this work (i.e., review on proposed solutions to enhance GDPR compliance) but on identifying critical success factors of GDPR implementations (Teixeira, 2019). The identification of success factors and barriers to comply with GDPR can help organizations to be better prepared to achieve compliance, by prioritizing those factors while avoiding possible obstacles.

Following this result, which lack proper content to examine, the authors decided to perform a search, using the same terms, on Google search engine. This did not retrieve any scientific published review articles, but only related content from other sources, mainly from industry reports or organization news, which the authors found pertinent to relate as a means to compare with their presented work (section 2.3).

2.3 GDPR Insights – One Year Later

Directly from the “horse’s mouth”, the European Commission has published, a year later (June 2019), a report on the impact of GDPR application on data protection (European Commission, 2019).

The report concludes that most Member States have set up the necessary legal framework for personal data protection enforcement. On the whole, most businesses are on the way to developing compliance while citizens are becoming more aware of data protection rules, and their rights. GDPR is also having an impact at the International level, where data protection authorities are cooperating more closely within the European Data Protection Board. By the end of June 2019, the cooperation mechanism had managed 516 cross-border cases. As more countries across the world equip themselves with modern data protection rules, they use the EU data protection standard as a *reference point*.

However, only 20% of Europeans know which public authority is responsible for protecting their data and still a minority fully reads privacy statements online. This is mostly because they are unclear and difficult to understand, or just knowing there is a privacy policy available, is enough.

At about the same period of time, this article (Klammer, 2019) confirms some of these conclusions, probably fed by the same results published by the European Commission’s survey. However, it also recalls that the GDPR for individual consumers, has led to a great increase on privacy

policy email updates from companies, on a rushed attempt to comply. This was also followed by a constant stream of consent pop-ups and cookie banners that Europeans need to face every day, when they navigate on the web. Contrary to the EU survey conclusions on putting GDPR as a reference point as data protection legislation, for U.S. companies that do business abroad, the GDPR represents a constant struggle to refine their data protection policies. This report finishes with a relevant message, instead of waiting to see how these laws are enforced, businesses should take proactive steps in securing consumer’s data and assessing compliance with GDPR.

On this last note, this report (SMEUnited, 2019) gives examples on how EU SMEs invested in awareness and advice to ensure that they comply with GDPR during the two-year transition period, and the past first year. Despite these efforts, taken together with the European Commission and the national authorities, there are still many questions on the application and implementation of this legislation. The main issues needing clarification are: a) controller vs processor, b) what *processing at large scale* means, c) record keeping of processing activities, and d) the principle of accountability. There are also difficulties in appointing a Data Protection Officer. The main conclusion from this report is that GDPR is still very complex to interpret and may require huge investments, which are usually not proportionate to the size of the organisations. The final message is that measures should be taken into reducing SME’s high economical and resource burden, and focus should be on providing them with the much-needed support, instead of just fining them.

On a more technical note, which discusses crucial security and privacy issues, GDPR requirements that work well in theory raise, in the real settings, unintended consequences, which can be very harmful for personal data protection (Stapp, 2019). Examples include the fact that, for impersonation attacks, when an account gets hacked, the hacker can use the right of access to get all data from the stolen account. Similar problems can happen for the right of data portability. Also, in relation to the right to be forgotten, this is applied blindly to any personal data, making it possible for anyone with a bad track record to hide, or send to oblivion, his/her problematic past record from the general public, which can become a public safety risk.

And finally, from all these analyses and discussions, researchers fear that scientific research can be hugely affected as GDPR can make harder for data to be shared across borders, or even outside their

original context. More so, if data belongs to special categories, such as health or genetics, where most breakthroughs are made because of big teams of researchers, who work on similar problems together, around the world, constantly share their work and outcomes.

So, in practice, what has been improved and innovated in terms of technologies and solutions to help attaining GDPR's compliance? This work focuses on answering this question.

3 METHODS

The literature review comprised papers written in English, published from 2018 until November 2019, which focus on GDPR solutions, tackling requirements on the way to improve GDPR compliance. For short, works that presented the applicability of the legislation in practice or in simulated use cases. Exclusion criteria included:

- Other similar reviews (although these could help in the justification and support for this work);
- Papers just describing and/or discussing GDPR requirements;
- Papers evaluating GDPR compliance in specific domains;
- Not free available full-texts;
- Papers not published in conferences or journals, such as editorials, letters and others.

Papers were searched in three different online research databases, namely: IEEE Xplorer, ACM and SCOPUS. Two different queries with the terms "GDPR" and "General Data Protection Regulation" were applied within the three database search engines. These terms were generic enough to include most works that mention or focus on GDPR.

The first part of the review was to analyse titles and abstracts of the list of articles retrieved by the two queries, according to the inclusion/exclusion criteria, as well as identify the repeated papers among different databases (Table 1).

A total of 51 papers was selected at this stage.

Table 1: Number of papers obtained from the search queries and selected from the three research databases, after titles and abstracts' review.

Research Database	"GDPR"	"General Data Protection Regulation"	Papers Selected
Xplorer	153	103	24
ACM	87	48	15
SCOPUS	820	639	12
TOTAL	1060	790	51

The second part of the review included reading the full available text, to verify again the inclusion/exclusion criteria, and extract main data points to be further analysed.

These data points comprised the:

- Year of publication
- Objective of the work
- Proposed GDPR solution
- Focused GDPR requirements
- Methodology
- Evaluation/proof of concept
- Sample characteristics (if applicable)
- Application to a real setting (Y/N)
- Main results/recommendations

Two papers were excluded at this second stage since they did not present GDPR solutions but were only assessing its applicability in terms of GDPR requirements, such as consent and portability.

The final data extraction for analysis was performed in a total of 49 papers.

The main methodology steps for the performed literature review are presented in Figure 1.

To notice that, since SCOPUS' database returned a very high number of papers from the search queries, the authors only perused the first 200 from each list of results. This can be justified by the fact that the returned list is shown according to the degree of relevance and relation to the search query terms, therefore, the list of 200 results will have the articles closely connected to the subject at hand. Obviously, this is according to the classification of relevance attributed by the SCOPUS database.

4 RESULTS

The final analysed sample comprised a total of 49 papers, 24 were published in the year of 2018 while

the remaining 25 in the year of 2019 (until November) (Figure 2).

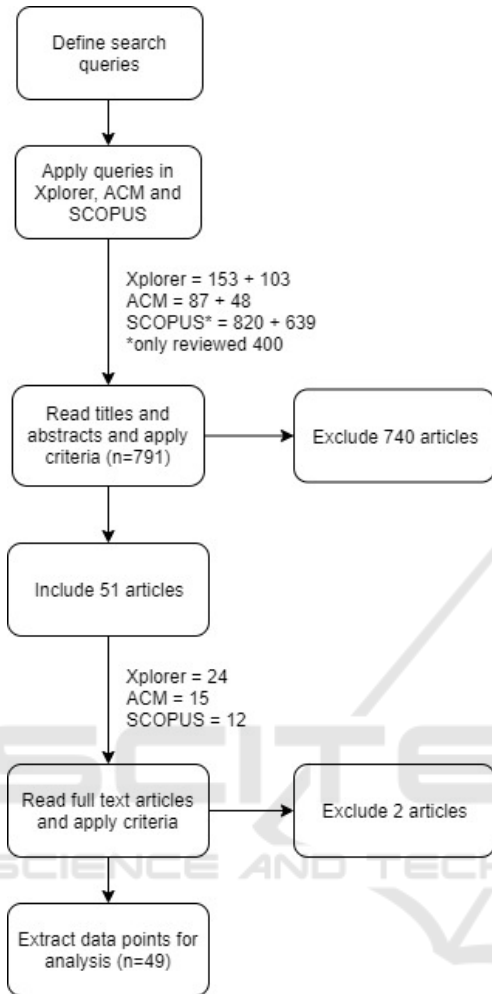


Figure 1: Flowchart of the methods used to perform the literature review.

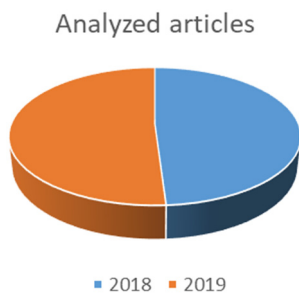


Figure 2: Number of analysed papers per year (n=24 in 2018 and n=25 in 2019).

Table 2 presents the number of papers that tackle specific GDPR requirements. The most common articles focus on researching GDPR requirements in

terms of consent (Art. 6 & 7), privacy by design and by default (Art. 25), as well as record of processing activities (Art. 30) and privacy impact assessments (Art. 35).

However, the biggest occurrence (with almost half of the sample) is the one that does not specify the type of GDPR requirements that are addressed, or mention several of them (usually more than 2 different requirements). These works are categorised in the “All compliance or not detailed” category.

Table 2: Number of papers that focus on each of the analysed GDPR articles/requirements (the number of occurrences can be greater than 100%, because one GDPR requirement can occur more than once in the same paper).

GDPR Article (s)	Number of occurrences	(%)
6 & 7	6	12
9	0	0
12	2	4
15, 16, 18 & 21	2	4
17	3	6
20	0	0
25	6	12
30	4	8
32	2	4
35	4	8
All compliance or not detailed	24	49

Figure 3 shows the most used technologies for data privacy in the GDPR solutions described in the analysed sample. These are: blockchain, homomorphic encryption, machine learning and PKI.

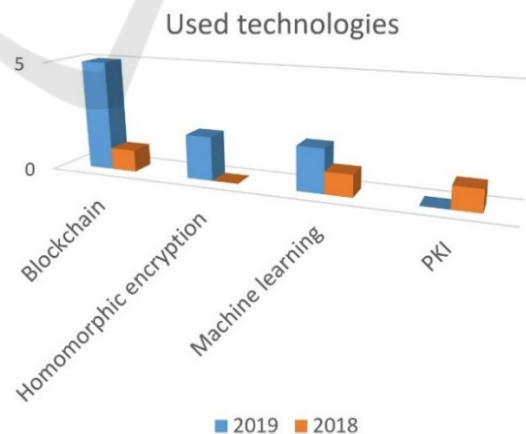


Figure 3: Most used technologies for the presented solutions, in the analysed sample.

The use of PKI technology related to GDPR solutions development has decreased from 2018 to 2019, while the other three technologies have taken a

different turn, and have recently increased. To notice that blockchain technologies were the ones with the highest increase in use.

In total, more than half of the analysed papers (n=29, 59%) refer that the presented solution was tested, validated or somehow evaluated in specific domain scenarios. Most frequent scenario for discussing the solution implementation/validation is healthcare (n=11(22%)), while the most used specific technology domain for tested application is the IoT (n=13(27%)), with a great increase in 2019 (Table 3). Moreover, risk assessment was also another area of research and test but in 2019, the focus seems to be decreasing, however, not all research for the year 2019 is published yet, and the numbers may still alter.

Table 3: Main areas of GDPR applicability research found in the literature review.

Areas	2018	2019	Total n (%)
IoT	4	9	13 (27)
Health	6	5	11 (22)
Risk assessment	6	3	9 (18)

Other areas of focus, although in much less number are: Privacy and data policies (3); Cloud-based architectures (2); transparent AI (1); CERTs (1); eHIFS - History Independent File Systems (1), Ontologies (1); LMS – Learning Management Systems (1); Ethics (1); Linked Widgets (1); SMEs (1); IDS – Intrusion Detection Systems (1); Intelligent Tutoring Systems (1); and PCI-DSS (Payment Card Industry Data Security Standard) (1).

Only 2 of the reviewed papers refer the test and implementation in real practice, in the domains of eLearning and a charity institution.

5 DISCUSSION

According to the obtained results, it may appear that is still early days for GDPR literature reviews, however, is it still early for GDPR knowledge and experiences to be more widely improved and shared? The answer is Yes, and No. Although there are not yet many scientific reviews on GDPR solutions, to generate irrefutable evidence on GDPR progress and improvements, there is enough evidence to declare that there are not enough appropriate solutions and support for businesses and organizations in general, to enhance their GDPR compliance.

Most scientific work on GDPR does not focus on providing and improving solutions to better face the GDPR key challenges and requirements. After a year

and a half, most research is trying to assess if GDPR has entered all the realms where there is the need to provide privacy and security of personal data. This is also relevant, and results are mostly not satisfactory. Examples such as these (Benjumea, 2019) (Muchagata, 2019) (Vlajic, 2018) are available and, since this was not the purpose of this work, no more references or similar examples, are provided here.

Regarding the sample analysed in this work, it was a balanced sample divided in the 2 years of analysis, however, the year of 2019 has not finished yet, and the number of papers could still increase, as publication procedures finish. Now, will this increase be in the number of solutions provided or again in works assessing the status of GDPR compliance? Probably more for the later, but the authors recommend the performance of regular reviews concerning both the issues, to make sure research can keep up and help in setting the best of GDPR requirements to all areas.

For the works that focus on solutions to improve GDPR compliance, these target commonly the areas of Consent management, Privacy by Design and Default, Recording of activities and Privacy Impact Assessments (PIA). The majority, however, deals with the subject of personal data protection without specifying GDPR requirements or target more than two of them. Maybe they try to deal with privacy in generic terms and the technologies that are mainly used (e.g., mostly cryptography based) only target those.

The main used technologies for the implementation of the proposed solutions are: blockchain, homomorphic encryption and machine learning. There seems to be a decrease in the use of PKIs from last year, maybe due to the fact that now blockchain and machine learning or artificial intelligence (AI) related algorithms, are technologies currently more used/adopted to develop security protocols, or to mine data. However, it is not clear if these technologies are the right ones to help on the path to GDPR compliance. AI can help improve system's correctness, but it makes the obtained decisions harder to explain or the right to be forgotten, even harder to be accomplished.

So, how will businesses prove and explain how security and protection is being provided for their clients' personal data? Can GDPR, in the end, contribute to avoid and stop enhancements for AI systems? Security and privacy regulations cannot be a reason or an obstacle to stop the development and advancement of new technologies.

In terms of the GDPR requirements that were not found in the reviewed works, it is surprising that

research is not tackling the issue of special categories (Art. 9) protection, as these are areas where huge amounts of resources are spent on research (e.g., bio engineering, medical informatics, mHealth, etc). Moreover, the main use case areas found in the presented review, and where solutions are tested, is healthcare. The need is there, so should be the focus on improved technology and solutions for GDPR compliance.

Similar results were found for the requirement on data portability (Art. 20). Not one solution was found to deal specifically with portability issues with IoT, where data needs to be privately shared and communicated over different types of devices, with different security and privacy requirements. Interestingly, again IoT is one of the main use cases where proposed solutions are tested, in the analysed sample. This can be explained by the fact that data portability seems easy enough to obtain but very hard to control. As already mentioned, data portability opens a hole/backdoor for hackers to easily exploit, once other attacks (e.g., impersonation/spoofing) are successful (Stapp, 2019).

The authors believe that more research work and resources need to be focused on the privacy of personal data in relation to mHealth and healthcare of IoT devices, which is certainly bound to grow immensely in the near future. Once there are appropriate solutions for this type of scenarios, these can be easily translated to most other scenarios, with the same or even less security and privacy requirements.

Although risk assessment does not seem at the moment one of the areas where use cases and testing scenarios are betting, the authors believe this will surely have a big focus in a near future, as PIAs are one of the main requirements for personal data privacy, anytime/anywhere, ideally, in a seamless manner. The authors have confidence in that, in order to make sure any business or organization that performs personal data processing is able to comply with GDPR, it needs to perform easy, simple but effective PIAs during the whole processing activities' lifecycle. *Before* – when processing needs are defined and set; *During* – when monitoring is needed to make sure (and to prove to the authorities) all is being done accordingly, and in the best way possible; and *After* – to understand what went wrong and why, as well as what went right and why, for future similar activities.

Again, more research efforts should be placed in PIA self-assessment tools, both for technical as well as administrative and human processes.

Besides the already specified areas of research that need more effort (e.g., PIA, data portability or

special categories), a more detailed analysis on what GDPR key requirements need more, or less, research and efforts, was not made in this work, but certainly needs to be taken into account, in future similar reviews.

In conclusion, it is true that only one year and a half has elapsed since GDPR came into force, however, there exists, in the reviewed literature, a very reduced amount of improved solutions, tested in real settings. This is not a good omen. More tests and experiences should be made in real settings, with real users so that solutions can be improved and be useful and, more importantly, be reproduced over similar scenarios. This will allow easier learning and adoption, as well as providing appropriate feedback to responsible legislators where adaptations to the regulation itself can, and probably should, be made along the way.

Limitations. The limitations of this work comprise the reduced sample for this specific literature review, where most of the works found on GDPR focus on: a) the analysis and recommendations of the requirements to comply with the regulation; or b) current assessments if GDPR is being properly adopted, in various areas.

Another limitation was the short time period of the review, one year and a half, however, it was already explained that this type of issue is relevant enough to be regularly assessed. Only this way can appropriate feedback be given to the responsible authorities and developers and researchers in general, for them to trust on a correction or continuation of the path they are taking.

This review was performed by only one researcher but due to the small sample size found, the authors do not consider there would be the need for an agreement rating on the included articles for analysis.

Another limitation was the lack of examination whether the topics that got funded by funding mechanisms such as the EU Commission or private companies, are oriented to specific GDPR requirements, or balanced to all of them.

6 CONCLUSIONS

The aim of this work is to make a literature review on proposed and introduced solutions to help and improve GDPR compliance, on its different requirements and challenges.

Although the sample found was not big and the analysed period of time is short, several

recommendations are drawn which can help current researchers and developers in the areas of security and privacy to rethink their direction.

In summary, there is the need for more novel solutions, tests and evaluations in real settings, to be easier/faster the achievement of knowledge to be shared to whom and where it is required. Furthermore, solutions on privacy cannot only be developed using cryptography or similar technologies. Solutions must integrate socio-technical components that can face in a more comprehensive and unabridged way, and not in isolation, complete infrastructures, activities and processes. Encrypted communication channels and data storage is not enough when social engineering and human related breaches are still light years away from being solved, which can negatively influence how personal data is protected.

Future work includes the regular performance of these reviews regarding the needs for technologies and solutions to comply with GDPR. This first review will aid in this direction, since future reviews can be done incrementally to this one, as they already have this work, as a reference for comparison.

ACKNOWLEDGEMENTS

This work is supported by TagUBig - Taming Your Big Data (IF/00693/2015) from Researcher FCT Program funded by National Funds through FCT (Fundação para a Ciência e Tecnologia).

REFERENCES

- Alizadeh, F., Jakobi, T., Boldt, J., Stevens, G., 2019. GDPR-Reality Check on the Right to Access Data: Claiming and Investigating Personally Identifiable Data from Companies. In *Proceedings of Mensch und Computer 2019 (MuC'19)*, 811-814.
- Benjumea, J., Dorrnzoro, E., Roperio, J., Rivera-Romero, O., Carrasco, A., 2019. Privacy in Mobile Health Applications for Breast Cancer Patients. In *IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS)*, pp. 634-639.
- Breitbarth, P., 2019. The impact of GDPR one year on. *Network Security*, Volume 2019, Issue 7, Pages 11-13.
- European Commission, 2019. General Data Protection Regulation shows results but work needs to continue. Press Release. Available online: <https://gdpr-info.eu>. Accessed: 27/12/2019.
- GDPR, 2019. *General Data Protection Regulation*. Available online: <https://gdpr-info.eu>. Accessed: 27/12/2019.
- Klammer, S., 2019. The GDPR: A year in review. Technology Law Source: Mapping the evolving legal landscape. *Porterwright*. Available online: <https://www.technologylawsource.com/2019/07/article/s/privacy-1/the-gdpr-a-year-in-review/>. Accessed: 27/12/2019.
- Momen, N., Hatamian, M., Fritsch, L., 2019. Did App Privacy Improve After the GDPR?. In *IEEE Security & Privacy*, vol. 17, no. 06, pp. 10-20.
- Muchagata, J., Ferreira, A., 2019. Mobile Apps for People with Dementia: Are They Compliant with the General Data Protection Regulation (GDPR)? In *Proceedings of the 12th International Joint Conference on Biomedical Engineering Systems and Technologies HEALTHINF2019*, Volume 2, 68-77.
- SMEUnited, 2019. SMEs Say GDPR Needs Reality Check. Available online: <https://smeunited.eu/news/smes-say-gdpr-needs-reality-check>. Accessed: 27/12/2019.
- Stapp, A., 2019. GDPR After One Year: Costs and Unintended Consequences. *Truth on the market*. Available online: <https://truthonthemarket.com/2019/05/24/gdpr-after-one-year-costs-and-unintended-consequences/>. Accessed: 27/12/2019.
- Teixeira, G., Mira da Silva, M., Pereira, R., 2019. The critical success factors of GDPR implementation: a systematic literature review. *Digital Policy, Regulation and Governance*, Vol. 21 No. 4, pp. 402-418.
- Vlajic, N., Masri, M., Riva, G., Barry, M., Doran, D., 2018. Online Tracking of Kids and Teens by Means of Invisible Images: COPPA vs. GDPR. In *Proceedings of the 2nd International Workshop on Multimedia Privacy and Security (MPS '18)*, 96-103.