A structured approach to GDPR compliance

Antonio Capodieci[a]**, Luca Mainetti**[a]

[a] Università del Salento, Lecce, Italy. {antonio.capodieci, luca.mainetti}@unisalento.it

Abstract European General Data Protection Regulation (GDPR, EU 2016/679), adopted from the European Parliament, has profoundly changed the legislative approach to the protection of personal data by the European Union. The GDPR requires organizations to make profound changes. Organizations have to move from an approach based to the adoption of minimum-security measures, provided by the EU Directive of 1994, to a proactive approach based on accountability. Organizations have to adopt systems of verification and continuous improvement and adopt principles such as privacy by design and privacy by default. Privacy by design calls for privacy to be taken into account throughout the whole engineering process. The adoption of GDPR, by an organization, raises the main question of how to audit the organization's adherence. This paper proposes a structured approach, based on business process modeling, to support compliance with the GDPR. We have come up with an approach that has to identify the most important key point (s) for GDPR compliance.

1 Introduction

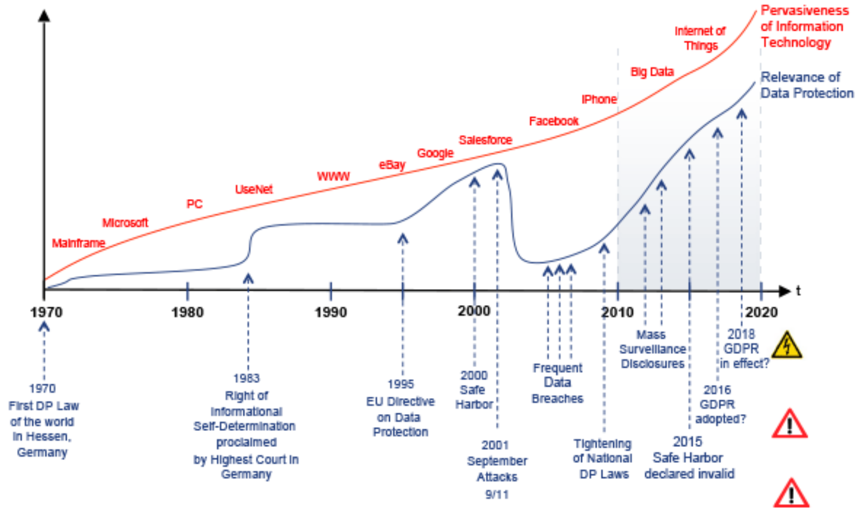
On 24 October 1995, the European Data Protection Directive (officially: Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data) was created as an essential element of EU privacy and human rights law. The directive required EU member states to implement the corresponding provisions in national law by 24 October 1998.

The Charter Of Fundamental Rights Of The European Union, approved in 2007, (2007/C 303/01) at the Article 8 - Protection of personal data, assert that "Everyone has the right to the protection of personal data concerning him or her." and "Such data must be processed fairly for specified purposes and on the basis of the consent of the person concerned or some other legitimate basis laid down by law. Everyone has the right of access to data which has been collected concerning him or her, and the right to have it rectified." The subsection 3 say "Compliance with these rules shall be subject to control by an independent authority."

As we can see in

Figure 1 [1], the panorama of “internet technologies and services” is completely changed from 1994 and the normative prescription became inadequate to protect personal data.

Figure 1 Internet Services & Regulations Evolution



Finally, starting from 2016, the European Union, basing principles established in the Charter Of Fundamental Rights has completely revolutionized the regulatory framework regarding the protection of personal data. Several legislative provisions (regulations and directives) have been issued and others are in the process of being issued.

EU introduces a common legal framework for all EU member states, with the aim of harmonizing their privacy principles and the application of these principles inside the Digital Single Market and overall to give, to individuals, more rights on their personal data.

Also important is that the standards apply to all companies that process data of European citizens regardless of the place in which these data are processed

The most well-known provision is the General Data Protection Regulation, better known as the GDPR.

One of the biggest innovations introduced by GDPR is self-assessment of digital risk and modulation of the required actions on the basis of an impact assessment analysis that includes specific measures to safeguard the data subject’s human dignity and fundamental rights.

In this context that organisations that collecting and processing personal data must be explicit about the motivation for data collection, who has access to data, and how, when and how many times the data will be used.

For an organization, to be GDPR compliant, is important to well explicit, and obviously, also to well know, how personal data are managed (the business process) and who deals with personal data (the actors in the process). A self-evaluation necessarily includes to understand all the business processes that are implemented and to identify who manage personal.

In the context of Computer Science, there are a variety of languages and methodologies and tools where the concept of a process takes centre stage. For example, DECLARE [2], DCR Graphs [3], State Charts [4], UML [5]–[8], GSM [9], CMMN [10], Business Process Modelling and Notation (BPMN) [11]. These instruments include details of both data collection and data use, but, all of them were developed before to the introduction of the GDPR, and the analysis phase does not preview to gather all the information necessary to ensure the compliance with GDPR.

Business Process Management Notation (BPMN) provides a standard, easy-to-read way to define and analyse business processes. BPMN creates a standardised bridge for the gap between the design of a business process and its implementation [12]. In the context of Data Protection framework Law, we believe that BPMN could be very useful and interesting.

2 The GDPR

The GDPR standardize legislation on the management of personal data throughout the European Union.

The Article 25 *Data protection by design and by default* requires that the protection of personal data be placed at the center of attention of the data controllers both in the planning and organization of services and in the stage of moedeling of IT systems. The subsection 1. says "*Taking into account the state of the art […] the controller shall, both at the time of the determination of the means for processing and at the time of the processing itself, implement appropriate technical and organisational measures, […], which are designed to implement data-protection principles, […], in an effective manner and to integrate the necessary safeguards into the processing in order to meet the requirements of this Regulation and protect the rights of data subjects*." In the subsection2. *The controller shall implement appropriate technical and organisational measures for ensuring that, by default, only personal data which are necessary for each specific purpose of the processing are processed. […].*

The records of processing activities [9, Article 30] are the main elements in the accountability of the owner, as they are useful in the recognition and evaluation of the treatments carried out and also in the risk analysis and proper planning of treatments. The register must contain at least the following information:(i)the name and contact details of the data controller; (ii)the purposes of the processing, distinguished by types of treatment; (iii)a description of the categories of data subjects (e.g. customers, suppliers, employees) and the categories of personal data (e.g. personal data, health data); (iiii)the categories of recipients (even by category only) to whom the personal data have been or will be communicated; (iiiii) the latest deadlines for the cancellation of the different categories of data; (iiiii) a general description of the technical and organisational security measures referred to in Article 32.

In general, GDPR are statements about how an organisation collects, processes, and more generally manages the personal data of individuals.

3 Motivation

To guarantee privacy by design during software development we can perform auditing and compliance checking, in ex-ante phase, and the detection of violations (ex-post phase) when they occur [13].

One key difficulty is that mainstream programming technologies do not allow software engineering to model the life cycle of the information; we don't have a view of the complete history or life cycle of the data [12]. We don't have a systematic overview of the processes and actors that deal with personal data.

But as author says in [14]–[16] a systematic approach, based on Business Process Management Analysis, is very useful to adapt business practices to emerging organizational forms.

So, we have hypothesized that a model based on BPMN can provide a suitable basis to support the Privacy By Design approach.

Moreover, BPMN could be useful to define records of processing activities that include all the information required by the GDPR, and can provide a solid support for auditing and compliance with GDPR, since BPMN allows an orientation to the business analysis domain.

4 Related Work

The state of the art presents several studies where BPMN meet security and privacy aspects. In[17] privacy concerns are captured by annotating the BPMN model. Brucker in [18], extend BPMN with access control, separation of duty, binding of duty and need to know principles.

In [19], BPMN is extended with information assurance and security modeling capabilities. Altuhhov et al in [20], aligned BPMN to the domain model of security risk management. In [21], Privacy Enhancing Technologies (PETs) are applied to enforce privacy requirements and support the analysis of private data leakage. A query language for representing security policies and a query engine that enables checking, is presented in Salnitri et al. [22].

Moreover, some works are related to the definition of extensions of BPMN to represent cyber security requirements [23], [24]. In [25], Maines et al. investigate an approach to modeling security. Authors used BPMN choreography to model message exchanged and identity contract negotiation.

Some work, [26], [27][15,17], present specific BPMN security extensions for healthcare context. Authors introduce security elements for BPMN to evaluate the trustworthiness of participants based on a rating of enterprise assets and to express security intentions such as confidentiality or integrity on an abstract level [28].

Nowadays a limited number of works have studied the correlation between the GDPR and process management. In our work we propose to extend BPMN with meta-information, for each element of BPMN, that classifies the element in the context of the GDPR.

There are few works that deal with GDPR, but, from different points of view.

The authors of [29] present "PRONTO: Privacy Ontology for Legal Reasoning", which is a first draft of a legal ontology for the GDPR that aims to provide a model of legal knowledge regarding the privacy agents, data types, types of processing operations, rights and obligations involved. This work is very interesting, and we will use this as a basis for our work, as described above.

The authors of [30] propose an approach that identifies a purpose with a business process, and show how formal models of inter-process communication can be used to audit or even derive privacy policies. Although this approach is also very interesting, in our opinion it is not able to identify all the information necessary to support GDPR compliance.

Finally, to capture security requirements within business process modelling, it is useful to have a notation that is supported by a set of graphical concepts, allowing us to represent the security semantics [27].

5 Proposed Approach

In this paper we proposed a methodological approach based on the analysis of business processes, which allows to precisely extract the records of processing activities with the necessary attributes. The proposed method also allows to identify all data processors.

5.1 Extension of BPMN

Our approach is based on an extension of the BPMN, already published in[31].

We defined a set of meta-information, for each element of the BPMN design, that classifies the element in the context of the GDPR.

For each pool/process, our method indicates:

(i) whether the process deals with personal data;
(ii) the legal basis that authorises its execution;

(iii) the period of time for which the data are stored. Each activity is classified as to whether it concerns personal data and the type of data processed.

In order to avoid creating custom BPMN notation extensions, the "tagged values" field was used. Based on the core definition of the BPMN, the appearance and specification of certain elements and connectors were defined by tagged values.

The following tags were inserted in the pool element:

- GDPR: IsPersonalDataProcessing: a Boolean value (Yes/No) indicates whether the process involves personal data.
- GDPR: LegalBasis: contains references to the motivations for the execution of the process.
- GDPR: Duration: the period of time for which storage is expected.

The following tags were inserted in the task element:

- GDPR: PersonalData: a Boolean Value (Yes/No) indicates whether the activity involves personal data
- GDPR: TypeOfPersonalData: indicates the type of personal data processed (personal data, judicial, health data, political and religious opinions, biometric).

In

Figure 2, already publish [31], in we can observe an application of our approach to the process of a vacation request.

First, we enter the expected value for the pool element.

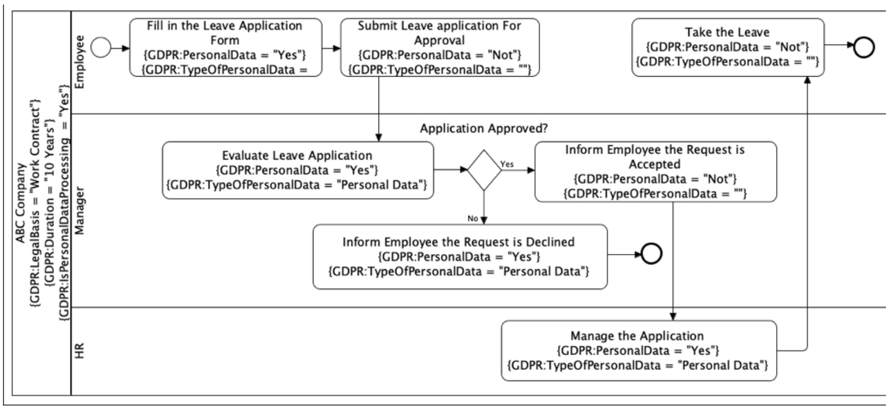
- GDPR:IsPersonalDataProcessing = Yes
- GDPR:LegalBasis= Work contract
- GDPR:Duration= One year

We can observe that the vacation request process involves records of processing activities and is related to the management of personal data.

After the pool, we analysed the tasks in each lane.

For each task, we defined whether it requires the management of personal data (using the tag value “GDPR:PersonalData”), and what type of personal data are managed (using the tag value “GDPR :TypeOfPersonalData”).

Figure 2 BPMN Vacation Request with GDPR Annotation


ABC Company
{GDPR:LegalBasis = "Work Contract"}
{GDPR:Duration = "10 Years"}
{GDPR:IsPersonalDataProcessing = "Yes"}
Employee
Manager
HR
Fill in the Leave Application Form {GDPR:PersonalData = "Yes"} {GDPR:TypeOfPersonalData =
Submit Leave application For Approval {GDPR:PersonalData = "Not"} {GDPR:TypeOfPersonalData = ""}
Take the Leave {GDPR:PersonalData = "Not"} {GDPR:TypeOfPersonalData = ""}
Application Approved?
Evaluate Leave Application {GDPR:PersonalData = "Yes"} {GDPR:TypeOfPersonalData = "Personal Data"}
Yes
No
Inform Employee the Request is Accepted {GDPR:PersonalData = "Not"} {GDPR:TypeOfPersonalData = ""}
Inform Employee the Request is Declined {GDPR:PersonalData = "Yes"} {GDPR:TypeOfPersonalData = "Personal Data"}
Manage the Application {GDPR:PersonalData = "Yes"} {GDPR:TypeOfPersonalData = "Personal Data"}

5.2 Proposed approach: BPMN-GDPR-ENHANCED

As we can see in Figure 3 our approach is based on 5 steps

Step1 - **Business Process Analysis**: The science has demonstrated the validity of Business Process Analysis to understand, with the necessary level of details, how organizations manage information and data. Then it's easy do hypnotize the application of BPA to discovery the process that manage personal data. BPMN is a consolidated tool to model Business Process.

Step2 - **Modeling using BPMN-GDPR Extended**: To discovery how organization manage personal data and to define the processor of such type of data, and to model activity (process) and actor (data processor) with the necessary information required form GDPR, we propose to adopt an extension of BPMN, see previous paragraph. In this way we

Step3 - **BPMN to XML**: In this stage we export the BPMN models of our organization, realized as showed in previous steps, in the XML format. In this way we have a machine-readable model of Business Process and Actors

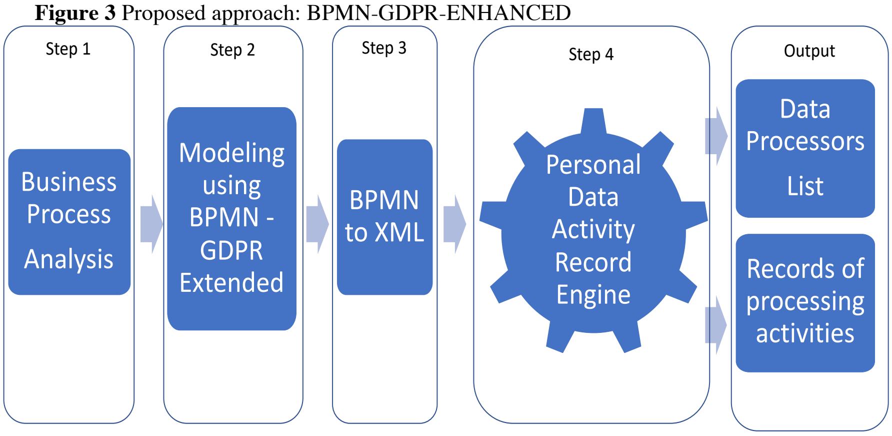
Step4 - Personal Data Activity Record Engine (PDARC): The personal Data Activity Record Engine is a tool able to query the XML model, coming from Step 2. PDARC is also able to extract all the process, and all the actors that manage personal data.

This is possible due to a classification based on the tagged values of the BPMN elements. The xml format of BPMN diagram, can be queried and a list of all the processes that have the tag "GDPR: PersonalDataProcessing" equal to true can be extracted.

The Output of this Step are tow element fundamental to demonstrate the compliance with GDPR:

- **Records of processing activities**
- **List of Data Processors**

Figure 3 Proposed approach: BPMN-GDPR-ENHANCED


Step 1
Business Process Analysis
Step 2
Modeling using BPMN - GDPR Extended
Step 3
BPMN to XML
Step 4
Personal Data Activity Record Engine
Output
Data Processors List
Records of processing activities

PDARC could export the "Records of processing activities" and the "List of Data Processors" in different "format", for example Word, EXCEL, XML, CVS, etc.. Some organization adopting specific application to manage the activity related to GDPR, then could be useful to import directly in this system the "Records of processing activities" and the "List of Data Processors" coming from PDARC.

At this moment PDARC is under development.

As we said before, we think, that this approach allows us to determine whether a process involves records of processing activities, with the necessary information to file the records, and to define the data processors.

7 Conclusion and Next Steps

The proposed methodological approach starting from the traditional business process analysis integrated with the contextual collection of information prescript from GDPR. From BPMN diagram, duly noted, we can extract the necessary information for the related to the records of processing activities (their legal basis, the duration of conservation, etc.). In the same way, it is possible to identify all actors who are data processors. As result, the preparation of the appointment of the employee as a data processor is facilitated.

In this way it is possible to quickly and completely draw up records of processing activities and to increase the accountability of an organisation with respect to the provisions of the GDPR. The ability to show that the records of processing activities and data processors were derived from the design and analysis of business processes can certainly increase the accountability of the organisation.

We are developing PDARC, the software to automatically extract the records of processing activities and the roles of the data processor by querying the BPMN modelling system.

We aim also to apply this approach in different context, we are working in the context of Local Public Administration. We are working also in the Health Care context where we deal with special categories of personal data that are object of a specific prescription, Article 9 of GDPR.

In the future we aim to create a tool that can be integrated with the run time company workflow engine that automatically generates records of processing activities and the appointments of the data processor that are aligned with the requirements of company processes.

Acknowledgment

This work was partially supported by the "EASYPAL" project no. Y95B457, funded by Innolabs Apulia Region (Italy) entitled "Sostegno alla creazione di soluzioni innovative finalizzate a specifici problemi di rilevanza sociale" under the POR Puglia FESR 2014-2020 research program sub-azione 1.4.B.

References

[1] E.-O. Wilhelm, "A brief history of the General Data Protection Regulation.".

[2] M. Pesic, H. Schonenberg, and W. M. P. Van Der Aalst, "DECLARE: Full support for loosely-structured processes," in *Proceedings - IEEE International Enterprise Distributed Object Computing Workshop, EDOC*, 2007.

[3] T. T. Hildebrandt and R. R. Mukkamala, "Declarative event-based workflow as distributed dynamic condition response graphs," *arXiv preprint arXiv:1110.4161*, 2011.

[4] D. Harel, M. Politi, and I. Books24x7, *Modeling reactive systems with statecharts*. 1998.

[5] A. M. Fernandez-Saez, D. Caivano, M. Genero, and M. R. V. Chaudron, "On the use of UML documentation in software maintenance: Results from a survey in industry," in *2015 ACM/IEEE 18th International Conference on Model Driven Engineering Languages and Systems, MODELS 2015 - Proceedings*, 2015, pp. 292–301.

[6] O. M. Group, "OMG Unified Modeling Language TM (OMG UML), Superstructure v.2.5," *InformatikSpektrum*, 2015.

[7] P. Ardimento, D. Caivano, M. Cimitile, and G. Visaggio, "Empirical investigation of the efficacy and efficiency of tools for transferring software engineering knowledge," *Journal of Information and Knowledge Management*, vol. 7, no. 3, pp. 197–207, 2008.

[8] S. España, N. Condori-Fernandez, A. González, and O. Pastor, "An empirical comparative evaluation of requirements engineering methods," *Journal of the Brazilian Computer Society*, vol. 16, no. 1, pp. 3–19, 2010.

[9] R. Hull *et al.*, “Introducing the guard-stage-milestone approach for specifying business entity lifecycles,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2011.

[10] Object Management Group, *Case Management Model and Notation (CMMN)*. 2013.

[11] Object Management Group, *Business Process Model and Notation (BPMN) Version 2.0*. 2011.

[12] M. Cremonini, E. Damiani, S. C. di Vimercati, P. Samarati, A. Corallo, and G. Elia, *Security, privacy, and trust in mobile systems and applications*. IGI Global, 2005.

[13] M. Enamul Kabir, H. Wang, and E. Bertino, “A conditional purpose-based access control model with dynamic roles,” *Expert Systems with Applications*, 2011.

[14] C. Ardito, U. Barchetti, A. Capodieci, A. Guido, and L. Mainetti, “Business Process Design Meets Business Practices Through Enterprise Patterns:,” *International Journal of e-Collaboration*, vol. 10, no. 1, pp. 57–73, 2014.

[15] U. Barchetti, A. Capodieci, A. L. Guido, and L. Mainetti, “Modelling Collaboration Processes through Design Patterns,” *Computing Informatics*, vol. 30, no. 1, pp. 113–135, 2011.

[16] A. Capodieci, L. Mainetti, and L. Alem, “An Innovative Approach to Digital Engineering Services Delivery: An Application in Maintenance,” in *2015 11th International Conference on Innovations in Information Technology (IIT) (IIT’15)*, Dubai, UAE, 2015, pp. 336–343.

[17] W. Labda, N. Mehandjiev, and P. Sampaio, “Modeling of Privacy-aware Business Processes in BPMN to Protect Personal Data,” in *Proceedings of the 29th Annual ACM Symposium on Applied Computing*, New York, NY, USA, 2014, pp. 1399–1405.

[18] A. D. Brucker, “Integrating Security Aspects into Business Process Models,” *it – Information Technology it – Information Technology*, vol. 55, no. 6, pp. 239–246, 2013.

[19] Y. Cherdantseva, J. Hilton, and O. Rana, “Towards SecureBPMN - Aligning BPMN with the Information Assurance and Security Domain,” in *Business Process Model and Notation*, 2012, pp. 107–115.

[20] O. Altuhhov, R. Matulevičius, and N. Ahmed, “An Extension of Business Process Model and Notation for Security Risk Management,” *International Journal of Information System Modeling and Design (IJISMD)*, vol. 4, no. 4, pp. 93–113, 2013.

[21] P. Pullonen, R. Matulevičius, and D. Bogdanov, “PE-BPMN: Privacy-Enhanced Business Process Model and Notation,” in *Business Process Management*, 2017, pp. 40–56.

[22] M. Salnitri, F. Dalpiaz, and P. Giorgini, “Designing secure business processes with SecBPMN,” *Softw Syst Model*, vol. 16, no. 3, pp. 737–757, Jul. 2017.

[23] M. E. A. Chergui and S. M. Benslimane, “A Valid BPMN Extension for Supporting Security Requirements Based on Cyber Security Ontology,” in *Model and Data Engineering*, 2018, pp. 219–232.

[24] C. L. Maines, D. Llewellyn-Jones, S. Tang, and B. Zhou, “A Cyber Security Ontology for BPMN-Security Extensions,” in *2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing*, 2015, pp. 1756–1763.

[25] C. L. Maines, B. Zhou, S. Tang, and Q. Shi, “Adding a Third Dimension to BPMN as a Means of Representing Cyber Security Requirements,” in *2016 9th International Conference on Developments in eSystems Engineering (DeSE)*, 2016, pp. 105–110.

[26] K. S. Sang and B. Zhou, “BPMN Security Extensions for Healthcare Process,” in *2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing*, 2015, pp. 2340–2345.

[27] A. Rodríguez, E. Fernández-Medina, and M. Piattini, “A BPMN extension for the modeling of security requirements in business processes,” *IEICE Transactions on Information and Systems*, 2007.

[28] M. Menzel, I. Thomas, and C. Meinel, “Security Requirements Specification in Service-Oriented Business Process Management,” in *2009 International Conference on Availability, Reliability and Security*, 2009, pp. 41–48.

[29] M. Palmirani, M. Martoni, A. Rossi, C. Bartolini, and L. Robaldo, “PrOnto: Privacy Ontology for Legal Reasoning,” in *Electronic Government and the Information Systems Perspective*, 2018, pp. 139–152.

[30] D. Basin, S. Debois, and T. Hildebrandt, “On Purpose and by Necessity: Compliance under the GDPR,” in *Financial Cryptography and Data Security (FC)*, 2018.

[31] A. Capodieci and L. Mainetti, “Business process awareness to support GDPR compliance,” in *In Proceedings of ICIST ’19*, Cairo.