BPMN-Enabled Data Protection and GDPR Compliance

Antonio Capodieci[1], Mimma De Carolis[2], Stefano Lisi[3], Luca Mainetti[1], Roberto Paiano[1] and Mariavittoria Ugirashebuj[4]

[1] *Università del Salento, Lecce, Italy*
[2] *Openwork s.r.l., Bari, Italy*
[3] *Politecnico di Bari, Bari, Italy*
[4] *Exprivia S.p.A. Molfetta (BA), Italy*

Abstract

The European Parliament adopted the European General Data Protection Regulation (GDPR, EU 2016/679) which revolutionized the legislative framework for personal data protection within the European Union. The GDPR mandates organizations to shift from a passive approach, relying on minimum security measures outlined in the 1994 EU Directive, to a proactive accountability-based approach. Organizations are expected to implement verification systems, foster continuous improvement, and follow principles such as privacy by design and privacy by default. The latter principle emphasizes incorporating privacy considerations throughout the entire engineering process. The challenge for organizations lies in effectively auditing their compliance with the GDPR. This study proposes a structured approach based on business process modeling to aid in GDPR compliance. It involves identifying crucial compliance points for the GDPR. A case study is presented where the method is applied to a purchase of a health insurance policy process in the context of the Secure Safe Apulia project.

Keywords: privacy by design, BPMN enhanced, business process GDPR compliant, data protection, GDPR, Process Modelling, BPMN, Privacy Requirements

1. Introduction

On October 24, 1995, the European Data Protection Directive (officially known as Directive 95/46/EC) was established as a critical component of EU privacy and human rights legislation. The directive mandated EU member states to incorporate its provisions into their national laws by October 24, 1998.

The Charter of Fundamental Rights of the European Union, ratified in 2007 (2007/C 303/01), enshrines the right to personal data protection in Article 8. This article states that "every individual has the right to protection of their personal data" and that such data must be processed in a fair and lawful manner, with the individual's consent or another legally recognized justification. Additionally, individuals have the right to access and rectify their data, with the compliance of these rules monitored by an independent authority.

As we can see in Figure 1, based on [1], the panorama of “internet technologies and services” has completely changed since 1994 and the normative prescription became inadequate to protect personal data.

In 2016, the European Union overhauled its regulatory framework for the protection of personal data, drawing from the principles outlined in the Charter of Fundamental Rights. The EU has issued several legislative provisions (regulations and directives) and others are in the process of being enacted. The goal of the common legal framework is to harmonize privacy principles across all EU member states and give individuals greater control over their personal data within the Digital Single Market.

IS-EUD 2023: 9th International Symposium on End-User Development, 6-8 June 2023, Cagliari, Italy

EMAIL: antonio.capodieci@unisalento.it (A. Capodieci) ; mimma.decarolis@openworkbpm.com (M. De Carolis); stefano.lisi@poliba.it (S Lisi); luca.mainetti@unisalento.it (L. Mainetti); roberto.paiano@unisalento.it (R. Paiano); mariavittoria.ugirashebuja@exprivia.com (M. Ugirashebuj)

In 2016, the European Union overhauled its regulatory framework for the protection of personal data, drawing from the principles outlined in the Charter of Fundamental Rights. The EU has issued

several legislative provisions (regulations and directives) and others are in the process of being enacted. The goal of the common legal framework is to harmonize privacy principles across all EU member states and give individuals greater control over their personal data within the Digital Single Market.

One of the key provisions is the General Data Protection Regulation (GDPR), which is widely known. The GDPR introduces the self-assessment of digital risk and mandates organizations to perform an impact assessment analysis to safeguard the dignity and fundamental rights of data subjects. Organizations must be transparent about their motivations for data collection, who has access to the data, and how, when, and how often the data will be used.

To be GDPR compliant, organizations must have a clear understanding of how personal data is managed (the business process) and who is involved in the process. This requires a self-evaluation that includes a thorough examination of all business processes and the identification of all actors involved in personal data management.

There are various languages, methodologies, and tools in computer science, such as DECLARE [2], DCR Graphs [3], State Charts [4], UML [5-8], GSM [9], CMMN [10], and Business Process Model and Notation (BPMN) [11], that focus on the concept of a process. However, these were developed prior to the introduction of the GDPR and do not gather all the information required for GDPR compliance.

Business Process Model and Notation (BPMN) provides a standardized and easily understandable way to define and analyze business processes. It creates a bridge between the design of a business process and its implementation. BPMN may be useful and valuable in the context of the data protection framework law.

2. The GDPR

The GDPR standardises legislation for the management of personal data throughout the European Union. Article 25 Data protection by design and by default requires that data controllers focus on the protection of personal data in both the planning and organisation of services, and in the stage of modelling of IT systems.

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 Subsection 2 it is required that 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 [12, 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); (iv) the categories of recipients (even by category only) to whom the personal data have been or will be communicated; (v) the latest deadlines for the cancellation of the different categories of data; (vi) 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. BPMN- Business Process Model And Notation

The Business Process Model and Notation (BPMN) is the leading standard for business process modeling. It is maintained by The Object Management Group (OMG) and supported by a wide range of vendors and employed by numerous organizations. The objective of BPMN is to support business process modeling for both technical users and business users, by providing notation that is intuitive to business users, yet able to represent complex process semantics. The BPMN 2.0 specification also provides execution semantics as well as mapping between the graphics of the notation and other execution languages, particularly Business Process Execution Language. BPMN is designed to be readily understandable by all business stakeholders. These include the business analysts who create and refine the processes, the technical developers responsible for implementing them, and the business managers who monitor and manage them. Consequently, BPMN serves as a common language, bridging the communication gap that frequently occurs between business process design and implementation.

4. BPMN- Business Process Model And Notation

5. Motivation

We propose that auditing and compliance checking using BPMN can help ensure privacy by design during software development [13]. BPMN has the potential to provide a systematic approach to support the privacy by design approach and ensure compliance with the General Data Protection Regulation (GDPR). Our hypothesis is that BPMN can provide a suitable basis for modeling the life cycle of information and systematically capturing the processes and actors involved in handling personal data.

Mainstream programming technologies do not provide a comprehensive view of the life cycle of data, making it difficult to track and manage privacy-sensitive information [14]. However, as suggested in previous research, a systematic approach based on business process management analysis can help organizations adapt their practices to emerging forms [15, 16, 17].

By using BPMN, we can create a standardized representation of business processes that includes all the information required by the GDPR and provides solid support for auditing and compliance. BPMN also allows for a business-focused approach to privacy and data protection. With BPMN, we can perform auditing and compliance checks in the pre-implementation phase, and detect violations in the post-implementation phase, to ensure ongoing compliance with privacy laws and regulations.

6. Related Work

The field of security and privacy requirements has been extensively studied and numerous studies have shown that BPMN can effectively meet these requirements. In [18], BPMN is used to capture

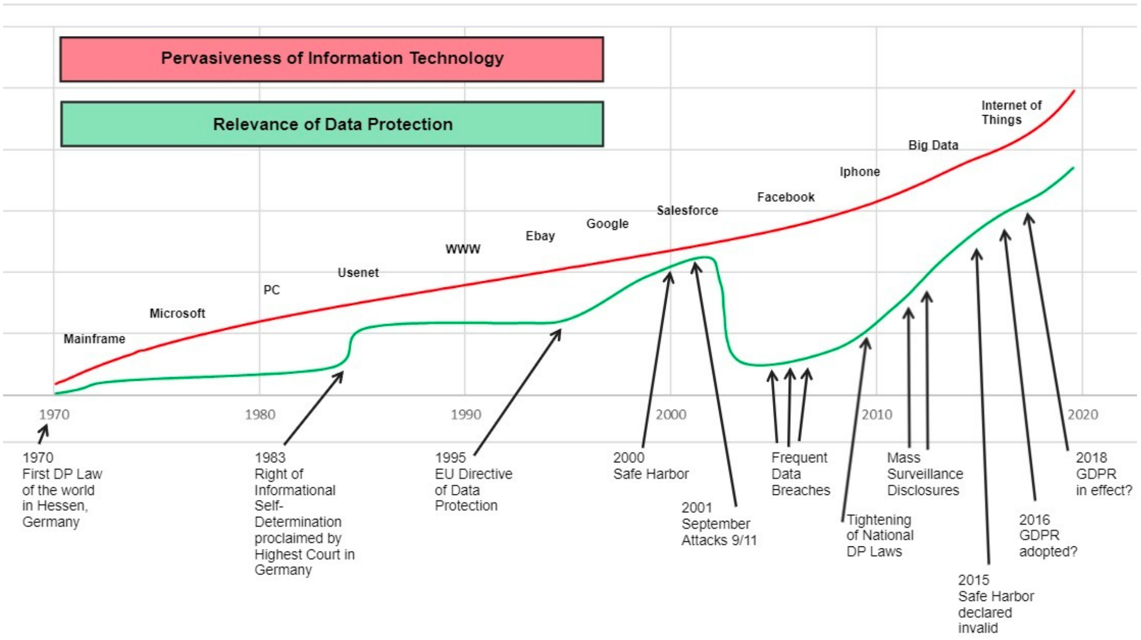

Figure 1 Internet Services & Regulations Evolution

privacy concerns through the process of annotating the BPMN model. Brucker, in [19], extends BPMN with access control, the separation of duty, binding of duty, and the need-to-know principles.

In [12], BPMN is extended with information assurance and security modelling capabilities. Altuhhov, in reference [20], aligns BPMN with the domain model of security risk management, while in reference [21], privacy enhancing technologies are applied to enforce privacy requirements and support the analysis of private data leakage. In [22], a query language for representing security policies and a query engine that enables checking are introduced.

In some studies, the authors define extensions of BPMN to meet cyber security requirements. For example, in [23] and [24], the authors investigate BPMN's ability to meet these requirements. Maines et al, in [25], study an approach to modeling security using BPMN choreography to model message exchange and identity contract negotiation.

There are also several studies that focus on BPMN security extensions in a healthcare context. For example, in [2, 21, 8, 3], the authors introduce security elements for BPMN to evaluate the trustworthiness of participants and to express security intentions, such as confidentiality or integrity, on an abstract level [26].

Despite the many studies that have been conducted on BPMN and security, there are only a limited number of studies that have examined the correlation between the GDPR and process management. In our study, we propose to extend BPMN with meta-information for each element of BPMN that classifies the element in the context of the GDPR.

In [27] the authors present “PRONTO: Privacy Ontology for Legal Reasoning”, which provides a legal ontology for the GDPR and models the privacy agents, data types, types of processing operations, and the rights and obligations involved. This work is of great interest and will serve as the basis for our work. In [28], the authors propose an approach that matches a “purpose” for data management, as defined in the GDPR, with a business process and demonstrates how formal models of inter-process communication can be used to audit or derive privacy policies. However, in our opinion, this approach does not capture all the information necessary to support GDPR compliance.

Finally, it is essential to have a notation that is supported by a set of graphical concepts in order to effectively capture security requirements within business process modeling. This allows us to represent the security semantics [29]. The use of BPMN with its well-established graphical concepts provides a suitable basis for this purpose. The health policy purchase process begins

7. Proposed Approach

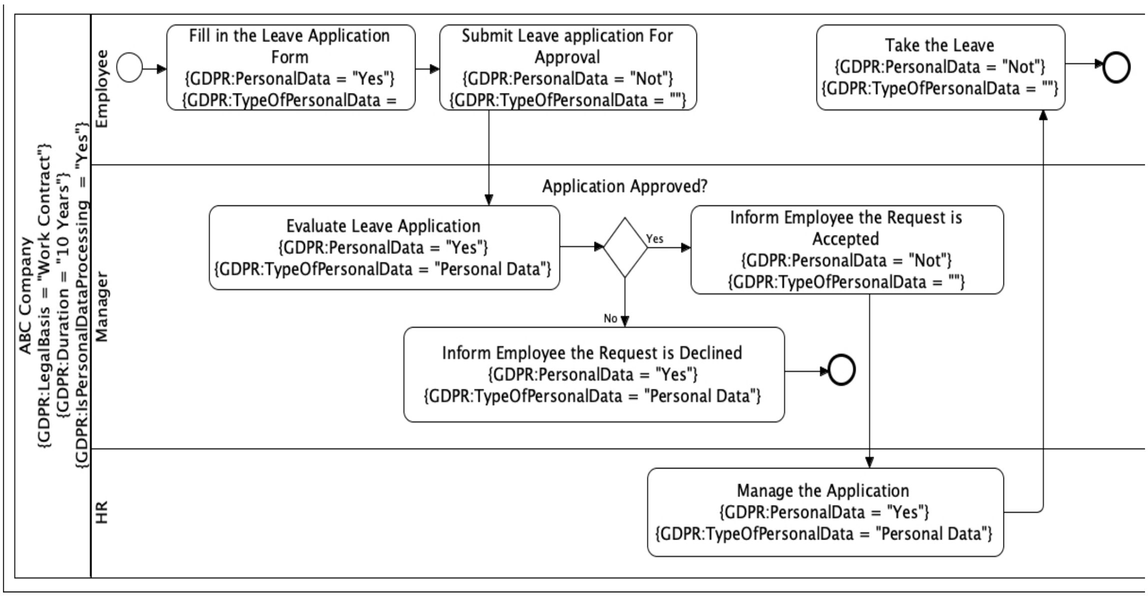
ABC Company
{GDPR:LegalBasis = "Work Contract"}
{GDPR:Duration = "10 Years"}
{GDPR:IsPersonalDataProcessing = "Yes"}
Employee
Manager
HR
Fill in the Leave Application Form {GDPR:PersonalData = "Yes"} {GDPR:TypeOfPersonalData =
Submit Leave application For Approval {GDPR:PersonalData = "Not"} {GDPR:TypeOfPersonalData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

Figure 2 BPMN Vacation Request with GDPR Annotation

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

7.1. Extension of BPMN

Our approach is based on an extension of the BPMN, already published in [30].

We have defined a set of meta-information for each element of the BPMN design, which 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 is stored. Each activity is classified as to whether it concerns personal data and the type of data processed.

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

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)
- GDPR risklevel: this tag identifies the scope of application, i.e. The level of risk present in the scope in which the data is used. A very high, high, medium and low risk is distinguished in relation to the parameters previously analysed
- GDPR RecipientData: finally, this tag identifies the users who, within the process, are authorized to access and process personal data. According to the art. 4 of the GDPR The recipient is "the natural or legal person, public authority, service or other organization that receives communication of personal data, whether or not it is a third party.

In Figure 2, already published [30], we show 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 see 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 is managed (using the tag value “GDPR :typeofpersonaldata”).

7.2. Case study: purchase of a health insurance policy

For reasons of brevity and readability we will describe the process in text mode and represent in BPMN one of the sub-processes in the TO-BE version. This is a real business process managed by our partner in a large research project.

The health insurance purchase process begins when the customer logs in to the company website.

The customer may not have an account. In this case it is necessary to activate the "on boarding procedure" for the creation of the new account and the delivery of the credentials.Once logged in to the site, the customer requests the policy and the company starts the card selection procedure (distinguishing between new and non-new users) and checks the validity of the card. Then start signing up for the policy by generating a series of codes required for registration.

We then proceed with the booking which gives rise to different situations. In particular, after having booked the visit, the date of the appointment is expected which can have three different outcomes, a positive outcome, when the customer shows up for the appointment, in this case the status of the appointment is first updated, the policy is activated and finally the payment is made which can generate the activation of the policy, when the payment is successful, or the non-activation of the policy, when there are problems with the payment. Also, you may have a negative outcome and, in this case, we proceed to charge you for the visit and update your policy status in relation to the payment. Finally, there may be a situation in which the user is absent and therefore the status of the policy and the appointment is updated.

As seen above, the health policy purchase process is divided into several sub-processes:

- On boarding procedure
- Paper selection procedure
- Policy signing procedure
- Appointment booking procedure
- Payment procedure

Referring to the sub-process related to the signing of the policy, which turns out to be the heart of the activity carried out by the main process, the company starts the subscription of the policy by generating a series of codes that the customer receives via SMS and which he will have to insert in the own web area. This phase, therefore, is characterized by a communication activity aimed at transforming unstructured information into structured information, using web 2.0 tools. In this sub-process the tags "gdprpersonaldata", "gdprtypeofpersonaldata", "gdprrisklevel" and "gdprrecipientdata" are used with reference to the activity of entering the codes carried out by the customer when he receives codes from the company and the identification of the outcome positive or negative of the signature with subsequent update of the policy status.

The figure 3 describe the sub-process of signing the insurance policy with GDPR Tag.approach.

8. Conclusion and Next Steps

Our proposed approach to privacy by design aims to integrate traditional business process analysis with the collection of contextual information as required by the General Data Protection Regulation (GDPR). The use of BPMN diagrams allows for the extraction of important information related to the records of processing activities, such as the legal basis for processing and the duration of data conservation. It also identifies all actors who are data processors, making it easier to appoint employees as such and prepare complete records of processing activities. This enhances an organization's accountability with respect to GDPR provisions.

Our goal is to provide a model integrated with a company's runtime workflow engine, in order to automatically generating applications inherently compliant with the GDPR requirements.

9. Acknowledgements

This work partially fulfills the research objectives of the Secure Safe Apulia project that was funded by the Apulia Region (Italy) under the 6ESURE5 grant agreement PO FESR 2014-2020. The authors thank the engineers of Eusoft Srl, Exprivia spa, Macnil Srl, Openwork Srl, and Sysman Srl Italian companies for their valuable collaboration during the prototype development.

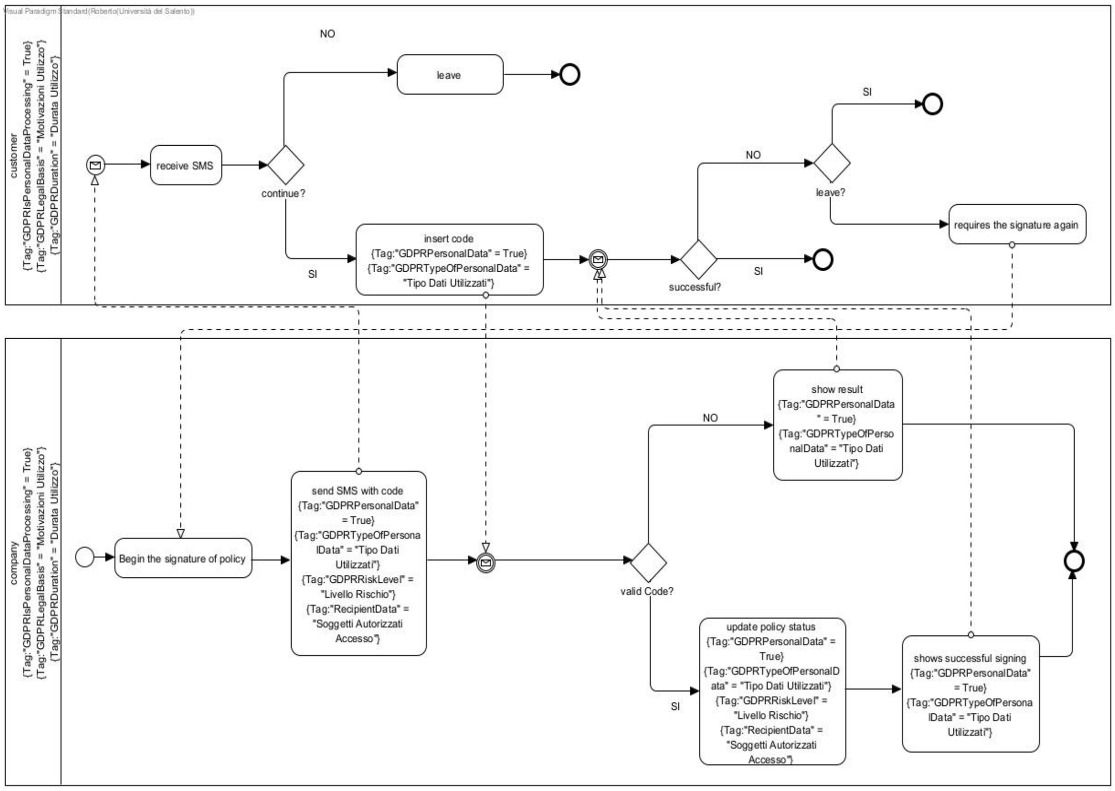
Visual Paradigm Standard(Roberto(Università del Salento))
customer
{Tag:"GDPRIsPersonalDataProcessing" = True}
{Tag:"GDPRLegalBasis" = "Motivazioni Utilizzo"}
{Tag:"GDPRDuration" = "Durata Utilizzo"}
receive SMS
continue?
NO
leave
SI
insert code {Tag:"GDPRPersonalData" = True} {Tag:"GDPRTypeOfPersonalData" = "Tipo Dati Utilizzati"}
successful?
NO
SI
leave?
SI
requires the signature again
company
{Tag:"GDPRIsPersonalDataProcessing" = True}
{Tag:"GDPRLegalBasis" = "Motivazioni Utilizzo"}
{Tag:"GDPRDuration" = "Durata Utilizzo"}
Begin the signature of policy
send SMS with code {Tag:"GDPRPersonalData" = True} {Tag:"GDPRTypeOfPersonalData" = "Tipo Dati Utilizzati"} {Tag:"GDPRRiskLevel" = "Livello Rischio"} {Tag:"RecipientData" = "Soggetti Autorizzati Accesso"}
valid Code?
NO
show result {Tag:"GDPRPersonalData" = True} {Tag:"GDPRTypeOfPersonalData" = "Tipo Dati Utilizzati"}
SI
update policy status {Tag:"GDPRPersonalData" = True} {Tag:"GDPRTypeOfPersonalData" = "Tipo Dati Utilizzati"} {Tag:"GDPRRiskLevel" = "Livello Rischio"} {Tag:"RecipientData" = "Soggetti Autorizzati Accesso"}
shows successful signing {Tag:"GDPRPersonalData" = True} {Tag:"GDPRTypeOfPersonalData" = "Tipo Dati Utilizzati"}

Figure 3 sub-process of signing the insurance policy with GDPR Annotation

10. References

[1] Wilhelm, E.-O.: A brief history of the General Data Protection Regulation, https://iapp.org/resources/article/a-brief-history-of-the-general-data-protection-regulation/.

[2] Pesic, M. Et al.: DECLARE: Full support for loosely-structured processes. In: Proceedings -

[3] Hildebrandt, T.T., Mukkamala, R.R.: Declarative event-based workflow as distributed dynamic condition response graphs. Arxiv preprint arxiv:1110.4161. (2011).

[4] Harel, D. Et al.: Modeling reactive systems with statecharts. (1998). Https://doi.org/10.1111/j.1462-2920.2012.02849.x.

[5] Ardimento, P. Et al.: Empirical investigation of the efficacy and efficiency of tools for transferring software engineering knowledge. Journal of Information and Knowledge Management. 7, 3, 197–207 (2008). Https://doi.org/10.1142/S0219649208002081.

[6] España, S. Et al.: An empirical comparative evaluation of requirements engineering methods. Journal of the Brazilian Computer Society. 16, 1, 3–19 (2010). Https://doi.org/10.1007/s13173-010-0003-5.

[7] Fernandez-Saez, A.M. et al.: 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. Pp. 292–301 IEEE Inc. (2015). Https://doi.org/10.1109/MODELS.2015.7338260.

[8] Group, O.M.: OMG Unified Modeling Language TM (OMG UML), Superstructure v.2.5. Informatikspektrum. (2015). Https://doi.org/10.1007/s002870050092.

[9] Hull, R. 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). Https://doi.org/10.1007/978-3-642-19589-1_1.

[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). Https://doi.org/10.1007/s11576-008-0096-z.

[12] Cherdantseva, Y. Et al.: Towards securebpmn - Aligning BPMN with the Information Assurance and Security Domain. In: Mendling, J. And Weidlich, M. (eds.) Business Process Model and Notation. Pp. 107–115 Springer Berlin Heidelberg (2012).

[13] Enamul Kabir, M. Et al.: A conditional purpose-based access control model with dynamic roles. Expert Systems with Applications. (2011). Https://doi.org/10.1016/j.eswa.2010.07.057.

[14] Cremonini, M. Et al.: Security, privacy, and trust in mobile systems and applications. IGI Global (2005). Https://doi.org/10.4018/978-1-59140-570-2.ch011.

[15] Ardito, C. Et al.: Business Process Design Meets Business Practices Through Enterprise Patterns: International Journal of e-Collaboration. 10, 1, 57–73 (2014). Https://doi.org/10.4018/ijec.2014010104.

[16] Barchetti, U. Et al.: Modelling Collaboration Processes through Design Patterns. Computing Informatics. 30, 1, 113–135 (2011).

[17] Capodieci, A. Et al.: 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). Pp. 336–343, Dubai, UAE (2015).

[18] Labda, W. Et al.: Modeling of Privacy-aware Business Processes in BPMN to Protect Personal Data. In: Proceedings of the 29th Annual ACM Symposium on Applied Computing. Pp. 1399–1405 ACM, New York, NY, USA (2014). Https://doi.org/10.1145/2554850.2555014.

[19] Brucker, A.D.: Integrating Security Aspects into Business Process Models. It – Information Technology it –55, 6, 239–246 (2013). Https://doi.org/10.1515/itit.2013.2004.

[20] Altuhhov, O. Et al.: An Extension of Business Process Model and Notation for Security Risk Management. International Journal of Information System Modeling and Design (IJISMD). 4, 4, 93–113 (2013). Https://doi.org/10.4018/ijismd.2013100105.

[21] Pullonen, P. Et al.: PE-BPMN: Privacy-Enhanced Business Process Model and Notation. In: Carmona, J. Et al. (eds.) Business Process Management. Pp. 40–56 Springer International Publishing (2017).

[22] Salnitri, M. Et al.: Designing secure business processes with secbpmn. Softw Syst Model. 16, 3, 737–757 (2017). Https://doi.org/10.1007/s10270-015-0499-4.

[23] Chergui, M.E.A., Benslimane, S.M.: A Valid BPMN Extension for Supporting Security Requirements Based on Cyber Security Ontology. In: Abdelwahed, E.H. et al. (eds.) Model and Data Engineering. Pp. 219–232 Springer International Publishing (2018).

[24] Maines, C.L. et al.: 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. Pp. 1756–1763 (2015). Https://doi.org/10.1109/CIT/IUCC/DASC/PICOM.2015.265.

[25] Maines, C.L. et al.: 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). Pp. 105–110 (2016). Https://doi.org/10.1109/dese.2016.69.

[26] Menzel, M. Et al.: Security Requirements Specification in Service-Oriented Business Process Management. In: 2009 International Conference on Availability, Reliability and Security. Pp. 41–48 (2009). Https://doi.org/10.1109/ARES.2009.90.

[27] Palmirani, M. Et al.: pronto: Privacy Ontology for Legal Reasoning. In: Kő, A. And Francesconi, E. (eds.) Electronic Government and the Information Systems Perspective. Pp. 139–152 Springer International Publishing (2018).

[28] Basin, D. Et al.: On Purpose and by Necessity: Compliance under the GDPR. In: Financial Cryptography and Data Security (FC). (2018).

[29] Rodríguez, A. Et al.: A BPMN extension for the modeling of security requirements in business processes. IEICE Transactions on Information and Systems. (2007). Https://doi.org/10.1093/ietisy/e90-d.4.745.

[30] Capodieci, A., Mainetti, L.: Business process awareness to support GDPR compliance. In: In Proceedings of ICIST '19. ACM, Cairo. Https://doi.org/10.1145/3361570.3361573. IEEE International Enterprise Distributed Object Computing Workshop, EDOC. (2007). Https://doi.org/10.1109/EDOC.2007.4384001.