Proceedings of the 35th European Safety and Reliability & the 33rd Society for Risk Analysis Europe Conference
Edited by Eirik Bjorheim Abrahamsen, Terje Aven, Frederic Bouder, Roger Flage, Marja Ylönen
 Published by Research Publishing, Singapore.
doi: 10.3850/978-981-94-3281-3_ESREL-SRA-E2025-P0510-cd

Ensuring Personal Data Compliance by integrating Legal Constraints into Digital Twin Design Methodology

Yann Lennon

Lab-STICC SHAKER, Université Bretagne Sud, yann.lennon@univ-ubs.fr

Nathalie Julien

Lab-STICC SHAKER, Université Bretagne Sud, nathalie.julien@univ-ubs.fr

Annabel Quin

Lab LEX, Université Bretagne Sud, annabel.quin@univ-ubs.fr

Advancements of digital twin technology have brought new use cases involving the presence of personal data. Considering the importance of this data, ensuring compliance towards the applicable legislation is mandatory. Through European provisions and especially the GDPR, this study aims to point out the various legal implications linked to the presence of personal data in the digital twin. Ensuring that processing operations comply with European regulations is not only a necessity, but also a key factor in risk management. In this paper, we propose an innovative approach where these constraints are formalized and integrated into an existing design methodology for digital twins. This study focuses on industrial digital twins in connection with maintenance operations.

When used for maintenance purposes, the use of a digital twin necessarily implies data sharing. These data flows must be considered and anticipated right from the design stage. It is necessary to conform to a framework that is comprehensible and accessible to its designers, to respect the binding rules linked to data processing in the architecture of the digital twin. To achieve a complete result towards compliance that is adapted to the specific use of the industrial digital twin, three steps will be discussed.

Firstly, the legal constraints must be identified directly from the use cases of digital twin technology. Secondly, the provisions of the GDPR are taken as a starting point to propose concrete measures that can be integrated when designing digital twins (record of processing activities, Data Protection Impact Assessment or DPIA, erasing policy). Finally, all these compliance operations will be brought together in a proposed methodology directly applicable to the design of digital twins. The design of this methodological framework represents an innovation in the field of digital twins : to ensure compliance with European standards, provide a framework for its design, and raise awareness among stakeholders designing digital twins.

Keywords: Digital Twin, Design Methodology, Compliance, Personal Data, European legislation

1. Introduction

Bringing together the realms of science and law poses a complex challenge, especially in the context of cutting-edge technologies such as digital twins. As the digital twin (DT) technology continues to gain popularity in the industrial world through Industry 4.0, it is also developing through multiple use cases like healthcare, territory replicas, etc. The rise of use cases for the digital twin technology has brought the presence of personal data for the purpose of assisting its users. However, the compliance operations regarding European regulations gets stricter when it comes to personal data.

Our study will focus on the European legislation applying to the European Union countries on personal data. Thus, national disparities can eventually occur. Each national legislation having its own specificity, it is necessary to get information on the national laws applying to the current project. The presence of personal data and

its use mentioned as "processing operations" is strictly governed by the General Data Protection Regulation [1]. The GDPR is the reference text for personal data protection, and this study aims to highlight its dispositions applied to the design and exploitation of a digital twin.

Designing digital twins requires a wide range of skills. In the face of such complexity, regulatory, ethical, and ecological aspects can be overlooked and dealt with too late. The aim of this study is therefore to integrate these notions as early as possible into the design phase of the digital twin, in a way that is consistent with existing methodologies for its technical aspects, aiming to avoid too much distance between technical and legal considerations [2]. These legal constraints must be integrated into a complete working framework; we have previously developed such a methodology based on generic architectures that still integrates human-system interactions as well as those between the digital twin and its environment [3-4].

The digital twin has specific characteristics, and the same applies to legal constraints. Although scientific literature exists for this general integration, it rarely deals specifically with digital twin technology [5] and the legal constraints are frequently tied as consequences of ethical constraints regarding the DT [6]. There are two main trends: firstly, there is the general and purely enumerative study of the various legal constraints from a general point of view [7-8-9]; then there are studies specific to extremely precise use cases, without even considering any typology of digital twins. As a result, there are numerous elements specific to each study that seem to limit it to a single, precise application such as smart cities [10-11], digital twins in the healthcare field [12-13-14], or precise aspects focused on specific variables [15-16].
The particularity of the study of legal constraints also relates to the law involved: the spatio-temporal framework of application is also important. Interesting studies are therefore limited to the incorporation of constraints specific to American law, which is not necessarily applicable on European territory [7].

This inclusion of legal constraints in the design of digital twins is preparatory work of general scope, with the aim of raising awareness of these constraints among all stakeholders, not just legal experts. In addition, this initial study makes it possible to discern the main steps involved in making the digital twin compliant in terms of personal data.

2. Scope and Scale

Considering that the DT is both a new and wide-ranging technology, establishing a typology of digital twins seems necessary when it comes to identifying specific legal constraints.

Initially, the term "digital twin" was defined in several different ways, leading to some confusion [5]. A definition is given by the ISO 23247 standard as: "*fit for purpose digital representation of an observable manufacturing element with synchronization between the element and its digital representation at an appropriate rate of synchronization*".

The breadth of the field of study applied to digital twins here concerns mainly industrial use with few specific legal constraints. Of course, digital twin technology is used in many different fields: health, military, administrative, industrial, etc. The elements discussed can be found in every case where personal data, as defined by the GDPR, is present. Here, the legal constraints evoked are as generalist as possible in order to be able to adapt this model to each use case. Nevertheless, each situation may entail specific legal implications as shown in figure 1. From four main types of digital twins, each one of them showcases certain legal constraints that would deserve a specific study. For the sake of simplicity, only the general considerations applicable to most projects, especially industrial ones, will be discussed here. Additionally, this methodology and these constraints have already been put to the test on current projects with an industrial dimension.

Figure 1: Identification of specific legal constraints for certain types of digital twins

3. Proposed Methodology

From studying the GDPR and the particularities related to the DT technology, we were able to identify recommendations and to propose a complete approach including legal compliance in a generic DT design methodology. Therefore, we designed a general scheme (Figure 2), representing the different DT lifecycle stages and the appropriate actions for compliance regarding GDPR. This general representation can afterward be detailed to create a risk assessment methodology for personal data. This scheme does not aim to represent an actionable algorithm. However, in the future, and after testing this methodology on case studies, it would be possible to present an algorithmic version of this scheme.

This representation is based on 6 steps included in 4 different phases of digital twin design and deployment. During the Configuration phase, every type of data involved in the digital twin must be identified and characterized as personal data, sensitive data, critical data, or regular data. After this identification, the different processing operations are described. From the diverse types of data and processing operations already existing, the risk assessment methodology (DPIA) is conducted in the Design Phase. The Deployment phase consists of the application of the solutions identified in the risk assessment method conducted previously. At last, the Exploitation phase represents the permanent attention required even if the technology is already deployed during the whole DT lifetime.

All those phases can be looped as many times as new data is introduced in the digital twin. So, every time a data is added, going through these steps will guarantee the DT is still compliant. The insertion of this methodology is part of a precise context. In fact, it integrates the 5A methodology, already in use, which is based on 5 main stages for the digital twin data: acquisition, aggregation, analysis, application, access [20] that has been deployed in various applications.

The compliance operations presented here will follow the order of these stages, which make up the life cycle of the digital twin.

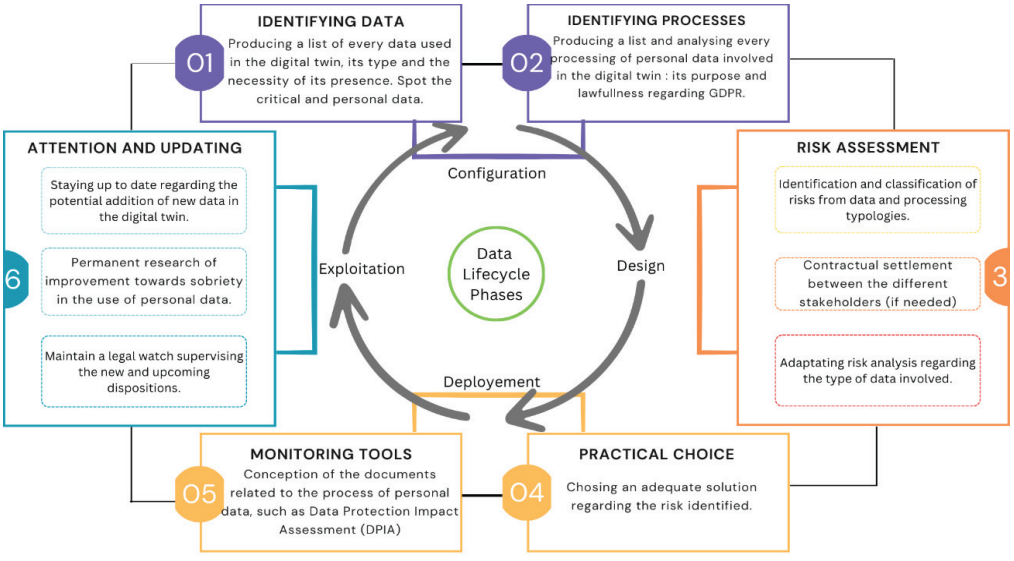

Figure 2: Personal data compliance process during the DT lifecycle

4. Data Qualification

4.1 Personal data

The use of personal data must be identified as soon as possible in the DT design in accordance with the criteria and definitions given by the GDPR: "*any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person*" (Article 4 from the GDPR).

From this definition, we can identify two distinct types of personal data: explicit and deductible. Those levels imply a need of vigilance and attention from the digital twin concepts and users. As well as the explicit personal data, every data traveling through the digital must be checked to make sure it does not refer to a living individual.

4.2 Sensitive data

If the personal data that feeds the digital twin is related to medical information, pathologies or patient information, the level of sensitivity is far superior from basic personal data. (Article 6 of the GDPR).

As opposed to personal data this specific information cannot be processed unless they respond to the conditions explicitly given by the GDPR. Amongst the different conditions described in the 9th article of the GDPR, we can mention the explicit and valid consent from the data subject and the processing's necessity to protect the vital interests of the data subject.
Both options are the most relevant regarding the digital twin applied to healthcare. However, the mindfulness regarding these possibilities is superior to regular personal data. We can mention this higher security level in the criteria forming the validity of the consent given from a patient but also to the storage time of every data. In addition to the conditions of processing operations

mentioned, every single rule applied to personal data is reinforced for these specific data [12-13], [17].

4.3 Critical data

The sensitivity of a data is not necessarily related to the fact that it's a personal data or not. Even if the GDPR and most of the legal dispositions are related to personal data, it is impossible to ignore the fact that some data could have critical importance without being related to a human person. A specific data used in the digital twin could be under the yoke of trade secrets.

Mentioning the existence of critical data is mandatory. It is crucial for the controller and the various developers to pay general attention to all the data required in the DT. This data, because of its importance, must also be processed, stored, and shared with specific requirements.

5. Data processing in digital twin design

The processing of personal data forces the compliance of the digital twin regarding the GDPR, with different obligations mentioned: the lawfulness, fairness, and transparency of processing personal data, the legitimate, specified, and explicit purposes of processing personal data, the respect of the data minimization principle limiting the processing to necessary purposes, accuracy and relevance of personal data, and appropriate and sufficient measures to ensure security of the personal data.
The lawfulness and the purpose of processing personal data are necessary criteria to ensure compliance within the use of the digital twin technology.

Regarding the lawfulness of processing, this criterion is explained by Article 6 of the GDPR. Considering the nature of the digital twin technology, we can already identify three common scenarios justifying the processing.

The **consent from the data subject** to process its data. The notion of consent is precisely defined by the GDPR as "*any freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her.*" (Article 4 of the GDPR). In the case of a digital twin, consent of the users should be asked prior to its operation in a way that ensures compliance regarding the GDPR. The continuation and performance of a contract to which the subject of data is a party.

The necessity for the "*purposes of the **legitimate interests** pursued by the controller or by a third party*" (Article 6 of the GDPR). This notion of legitimate interest is often the subject of debate [18] and must constantly be weighed against the rights and interests of the data subject. To take these nuances into account, the controller may consider several characteristic elements, such as the nature of the data, the characteristics of the data subject, and the implementation framework of the DT [19].

The **purposes of the processing** must be identified, especially its proportionality. Then, all the processing of personal data in the digital twin technology must be justified and necessary. Nonetheless, this operation could be problematic if the digital twin evolves and needs any new personal data. This new personal data will have to find its own purpose and be imperative.

6. Contractual settlements between stakeholders

The implementation and design of a DT within a structure necessarily entails the establishment of a contractual framework. This framework must enable the various stakeholders to ensure a sound implementation that respects everyone's wishes. Several crucial elements must be defined and specified contractually: scale, trade secrets and patent rights, data ownership, participating entities consent, and data storage and accessibility.

Defining the **scale** of implementation is necessary, as considering the broadness of use cases, a digital twin can concern multiple processes and implies different actors. In the case of an industrial digital twin monitoring different

processes operated by different stakeholders, multiple factors must be anticipated. All these constraints will have to be settled through contract negotiations. The awareness related to the use of data for every participating entity is required.

Concerning **trade secrets and patent rights**, it is obvious that the different participants in the digital twin, especially in a context of corporate relations, want to protect both their secrets and their own interests.

In the case of a single actor digital twin, the **data ownership** subject is not particularly relevant as it is internally stored. However, when it comes to multiple actor scenarios, this is one of the most important matters regarding the implementation of the digital twin. The implementation scenario of the technology must be as clear as possible. We can imagine a lot of several types of use cases.

Concerning **participating entities consent**, every actor must be able to refuse to participate in the digital twin, especially if the presence of personal data is involved. Additionally, they must be clearly informed of the different processing operations of their data.

Since multiple entities are involved, **data storage and accessibility** must be clarified during the contractual settlement. The type of data will influence the rules concerning this ownership and the possibilities opened to the different contractual parties. Collective ownership of data may also be one of the options considered to ensure equity between the digital twin's stakeholders. These collective ownership conditions will have to be determined through contractual negotiations.

7. Practical Consequences

These legal and contractual considerations oblige the DT designer to take concrete measures. These necessary measures have an impact on the progress of the project and the parties involved. They include the creation and maintenance of specific registers, the appointment of competent managers and the implementation of a data management policy.

7.1 Record of Processing Activities

Consequently, the presence of a controller implies a record of processing activities involving the presence of personal data. The article 30 of the GDPR clearly describes the different information this record must contain: the name and contact details of the controller and, where applicable, the joint controller, the controller's representative and the data protection officer; the purposes of the processing; a description of the categories of data subjects and of the categories of personal data; the categories of recipients to whom the personal data have been or will be disclosed including recipients in third countries or international organizations; where applicable, transfers of personal data to a third country or an international organization, including the identification of that third country or international organization and, in the case of transfers referred to in the second subparagraph of Article 49(1), the documentation of suitable safeguards; where possible, the envisaged time limits for erasure of the various categories of data; where possible, a general description of the technical and organizational security measures referred to in Article 32(1).

7.2 Nomination of a Controller

Additionally, the presence of personal data implies the designation of a controller, which is considered as "*the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data…*" (Article 4.7 of the GDPR). The presence of a nominated controller is mandatory.

According to this definition, a stakeholder in the digital twin must be designated as a controller. In fact, the principles previously mentioned must be monitored and (if necessary) proven to be respected by the processor. The controller has an obligation of accountability regarding the principles given by the GDPR (article 24 of the GDPR).

7.3 Data Protection Impact Assessment (DPIA)

The GDPR considers the use of new and innovative technologies. Although there is no

regulatory definition of what constitutes an innovative technology, we consider the DT to be one into this category.

This consideration brings into play a specific assessment required by the GDPR in its 30th article, the data protection impact assessment, strongly recommended when the risks involved in the personal data processing is high. However, the European regulation allows the different supervisory authorities to establish a list of the different scenarios requiring this assessment.

Consequently, the controller must check the different criteria given by the competent supervisory authority. Most of the different criteria could be identified in different uses of the digital twin technology, and the use of an innovative technology covers all the types of use cases.

Therefore, even if the DPIA is not absolutely required from the controller, it is highly recommended to be done during the design phase of a digital twin. [21]. Unlike the register of processing, the DPIA is a result of a defined risk assessment methodology. The DPIA methodology revolves around precise steps described by the 35th Article of the GDPR.

From this article, we can identify 4 mains steps: Explicit description of processing operations, Evaluation of the purpose, necessity, and proportionality of these processing operations, Identification of the risks related to these operations, and Description of the proposed solutions.

This 4-step methodology is close to other Risk assessment methodologies that already exist in other areas, typically the FMECA (Failure Mode and Effect Critical Analysis) in maintenance. The DPIA must be drafted in a reference document for the controller of the personal data processing operations. This methodology has already been successfully integrated in assisted tools, facilitating its implementation [22].

7.4 Erasing Policy for Personal Data

Identification of processing activities related to personal data must take into consideration the full life cycle of every data. Thus, the GDPR clearly describes the general aim every erasing policy should have personal data shall be "*kept in a form which permits identification of data subjects for no longer than is necessary for the purposes…".* Article 5, k) of the GDPR.

As we mentioned regarding the lawfulness of the processing operations, the principle of proportionality and sobriety is called again. Thus, the storage time of the personal data must be as limited as possible. To respect this principle of proportionality we can mention 2 different options: reducing the storage time for every data and showing a clear erasing method or creating a technical solution allowing to store the data without making it a personal data (anonymization process for example).

Therefore, the erasing policy of personal data for the digital twin technology requires attention both before the processing operations and during these operations. Taking into consideration the data lifecycle is necessary to ensure compliance [20].

8. Conclusion

Organizing compliance operations relating to the presence of personal data in the digital twin is essential in view of the development of this technology. Through this study, we have been able to demonstrate that at every stage in the life of the digital twin, there are specific operations to ensure the legal compliance of the twin. As these methods already exist for other uses, integrating them into a rapidly developing technology like the digital twin in a formalized way will ensure its safety, reliability and resilience for the future.

Bibliography

1. European Union. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 (General Data Protection Regulation). Official Journal of the European Union L119 (2016): 1–88.

2. Marchant, G.E. "The Growing Gap Between Emerging Technologies and the Law." In *The Growing Gap Between Emerging Technologies and Legal-Ethical Oversight*, edited by G. Marchant, B. Allenby,

and J. Herkert, 19–32. The International Library of Ethics, Law and Technology, vol. 7. Dordrecht: Springer, 2011.

3. Hamzaoui, M. A., and N. Julien. "Social Cyber-Physical Systems and Digital Twins Networks: A Perspective About the Future Digital Twin Ecosystems." *IFAC-PapersOnLine* 55, no. 10 (2022).

4. Hamzaoui, M. A., and N. Julien. "A Generic Deployment Methodology for Digital Twins – First Building Blocks." In *Handbook of Digital Twins*, 2024.

5. Kerckhove, D. "The Personal Digital Twin, Ethical Considerations." *Philosophical Transactions of the Royal Society A* 379, no. 2198 (2021).

6. Labadie, C., and C. Legner. "Personal Data Life Cycle for Data Protection." *Enterprise Modelling and Information Systems Architectures* 15, no. 9 (2020).

7. Zoltick, M.M., and J.B. Maisel. "Societal Impacts: Legal, Regulatory and Ethical Considerations for the Digital Twin." In *The Digital Twin*, edited by N. Crespi, A.T. Drobot, and R. Minerva, pages 825–841. Cham: Springer, 2023.

8. Janeček, V. "Ownership of Personal Data in the Internet of Things." *Computer Law & Security Review* 34, no. 5 (2018): 1039–1052.

9. Teller Marina. "Legal Aspects Related to Digital Twin." *Philosophical Transactions of the Royal Society A* 379, no. 20210023 (2021).

10. Bundin, M., A. Martynov, and E. Shireeva. "Legal Issues on the Use of 'Digital Twin' Technologies for Smart Cities." Communications in Computer and Information Science, vol. 1529. Cham: Springer, 2022.

11. Suffia, Gabriele. "Legal Issues of the Digital Twin Cities in the Current and Upcoming European Legislation: Can Digital Twin Cities Be Used to Respond to Urbanisation Problems?" In *Proceedings of the 15th International Conference on Theory and Practice of Electronic Governance (ICEGOV '22)*, 534–537. New York: ACM, 2022.

12. Erol, T., A.F. Mendi, and D. Doğan. "The Digital Twin Revolution in Healthcare." In *2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, 1–7. Istanbul, 2020.

13. Lal, A., J. Dang, C. Nabzdyk, O. Gajic, and V. Herasevich. "Regulatory Oversight and Ethical Concerns Surrounding Software as Medical Device (SaMD) and Digital Twin Technology in Healthcare." Annals of Translational Medicine 10, no. 18 (2022).

14. Drummond, D., and A. Coulet. "Technical, Ethical, Legal, and Societal Challenges With Digital Twin Systems for the Management of Chronic Diseases in Children and Young People." Journal of Medical Internet Research 24, no. 10 (2022).

15. Perez, G.C., and B. Korth. "Digital Twin for Legal Requirements in Production and Logistics Based on the Example of the Storage of Hazardous Substances." In 2020 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), 1093–1097. Singapore: IEEE, 2020.

16. Arda, Aslı. "A High-Tech Crystal Ball: The Digital Twin of a Ship and Data Issues in EU Law." Lloyd's Shipping and Trade Law 22, no. 5 (2023).

17. Chiruvella, V., and A.K. Guddati. "Ethical Issues in Patient Data Ownership." Interactive Journal of Medical Research 10, no. 2 (2021).

18. Ferretti, Federico. "Data Protection and the Legitimate Interest of Data Controllers: Much Ado About Nothing or the Winter of Rights?" Common Market Law Review 51, no. 3 (2014).

19. Doe, J., and R. Roe. "Legitimate Interest." In Elgar Encyclopedia of Law and Data Science, edited by A. Smith and B. Brown, 209–215. Cheltenham: Edward Elgar Publishing, 2024.

20. Lee, Yong, and Goo Yeon Lee. "Security Management Suitable for Lifecycle of Personal Information in Multi-User IoT Environment." Sensors 21, no. 22 (2021).

21. Article 29 Data Protection Working Party. Guidelines on Data Protection Impact Assessment (DPIA) and Determining Whether Processing is "Likely to Result in a High Risk" for the Purposes of Regulation 2016/679. 17/EN WP 248 rev.01. Adopted on April 4, 2017, revised October 4, 2017.

22. Dashti, S., and S. Ranise. "A Tool-Assisted Methodology for the Data Protection Impact Assessment." In Proceedings of the 16th International Joint Conference on e-Business and Telecommunications (ICETE 2019), 276–283. 2019.