Engineering a Digital Twin for Diagnosis and Treatment of Multiple Sclerosis

Giordano d'Aloisio
giordano.daloisio@graduate.univaq.it
DISIM Department
University of L'Aquila
L'Aquila, Italy

Alessandro Di Matteo
alessandro.dimatteo1@graduate.univaq.it
A^{2}VI-lab, DISIM Department
University of L'Aquila
L'Aquila, Italy

Alessia Cipriani
alessia.cipriani03@icatt.it
A^{2}VI-lab, Diagnostic Imaging, Oncologic Radiotherapy and Hematology Department
Sacro Cuore Catholic University of Rome
Rome, Italy

Daniele Lozzi
daniele.lozzi@graduate.univaq.it
A^{2}VI-lab, DISIM Department
University of L'Aquila
L'Aquila, Italy

Enrico Mattei
enrico.mattei@graduate.univaq.it
A^{2}VI-lab, DISIM Department
University of L'Aquila
L'Aquila, Italy

Gennaro Zanfardino
gennaro.zanfardino@graduate.univaq.it
DISIM Department
University of L'Aquila
L'Aquila, Italy

Antinisca Di Marco
antinisca.dimarco@univaq.it
DISIM Department
University of L'Aquila
L'Aquila, Italy

Giuseppe Placidi
giuseppe.placidi@univaq.it
MESVA Department
University of L'Aquila
L'Aquila, Italy

ABSTRACT

Multiple sclerosis (MS) is a complex, chronic, and heterogeneous disease of the central nervous system that affects 3 million people globally. The multifactorial nature of MS necessitates an adaptive and personalized approach to diagnosis, monitoring, and treatment. This paper proposes a novel Digital Twin for Multiple Sclerosis (DTMS) designed to integrate diverse data sources, including Magnetic resonance imaging (MRI), clinical biomarkers, and digital health metrics, into a unified predictive model. The DTMS aims to enhance the precision of MS management by providing real-time, individualized insights into disease progression and treatment efficacy. Through a federated learning approach, the DTMS leverages explainable AI to offer reliable and personalized therapeutic recommendations, ultimately striving to delay disability and improve patient outcomes. This comprehensive digital framework represents a significant advancement in the application of AI and digital twins in the field of neurology, promising a more tailored and effective management strategy for MS.

CCS CONCEPTS

• **Applied computing** → **Bioinformatics**; • **Computer systems organization** → *Data flow architectures.*

KEYWORDS

Multiple Sclerosis, Digital Twin, Data Harmonization, Bioinformatics.

ACM Reference Format:
Giordano d'Aloisio, Alessandro Di Matteo, Alessia Cipriani, Daniele Lozzi, Enrico Mattei, Gennaro Zanfardino, Antinisca Di Marco, and Giuseppe Placidi. 2024. Engineering a Digital Twin for Diagnosis and Treatment of Multiple Sclerosis. In *ACM/IEEE 27th International Conference on Model Driven Engineering Languages and Systems (MODELS Companion '24), September 22–27, 2024, Linz, Austria.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3652620.3688249

MODELS Companion '24, September 22–27, 2024, Linz, Austria

ACM ISBN 979-8-4007-0622-6/24/09
https://doi.org/10.1145/3652620.3688249

1 INTRODUCTION

Multiple sclerosis (MS) is a chronic, inflammatory, demyelinating, and neurodegenerative disease of the central nervous system (CNS), heterogeneous, multifactorial, immune-mediated, caused by complex interactions between genetic background and environmental factors [18, 41]. It affects about 3 million people in the world, a number significantly underestimated due to the lack of reliable data from developing countries [35]. It mainly affects young adults, with onset usually between 20 and 40 years of age, and it has one prevalence approximately 3 times higher in women than in men. The typical pathological feature of MS is the accumulation of demyelinating lesions that develop in both the white matter and the gray matter of the brain and spinal cord.

The clinical manifestations and course of MS are heterogeneous: in most patients, the symptoms are observed in reversible episodes of neurological deficits (known as relapses) that usually last days or weeks and characterize the initial stages of the disease (i.e., clinically isolated syndrome (CIS), and relapsing-remitting multiple sclerosis

(RRMS)). Over time, the development of permanent neurological deficits and the progression of clinical disability become prominent (secondary progressive multiple sclerosis (SPMS)). A minority of patients show a progressive disease course from the beginning, which is defined as primary progressive multiple sclerosis (PPMS). On average, almost 80% of RRMS patients skip to SPMS within 20 years of onset. Annually, it is estimated that 2-3% of patients with RRMS convert to SMSP [30]. However, there is significant individual variability in the risk of progression, which is often detected late and outside the time window in which the available therapies modifying the course of the disease are effective.

Therefore, the chronic, heterogeneous, and multifocal nature of MS, also called the "disease of a thousand faces", requires a diagnosis strategy, complex, ubiquitous, and differentiated, as well as adaptive, monitoring, and treatment. This strategy should be personalized and adapted to the patient's individual needs and disease course and be continuously adapted. The progression of the disease is schematically represented in Figure 1.

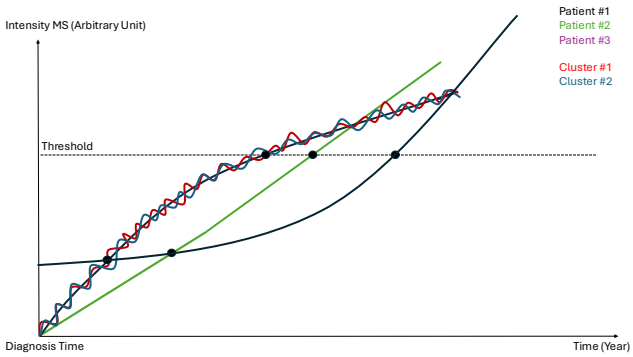

Figure 1: Exemplary progression of the MS according to classes of patients (smooth continuous lines), each generated by averaging the progression lines of single patients (high-frequency lines represented just around one class, to avoid confusion). Both types of curves are useful in MS: the class curves are required to indicate the current phenotype (the class) of the disease to which the specific patient currently allows; the patient's curve is required to calibrate specific treatments and to predict the future potential switch of a patient to another class (critical points are the crossing points among the class curves). The threshold line indicates the value above which the disability level is very high. The scope of the research is to use predictive markers to allow every patient to cross the threshold at the furthest point in time: in this optic, personal and cluster information have to be used jointly.

In recent years, research has focused on identifying reliable multiple-source molecular, diagnostic, and clinical biomarkers to improve clinical monitoring and predict the risk of progression of patients with early MS. The goal is to provide a system to allow patients to undergo disability furthest in time (Figure 1).

Magnetic resonance imaging (MRI) is the main technique for diagnosing and monitoring MS, thanks to its multimodal versatility in determining the load of demyelinating lesions in the CNS and disease activity (appearance of new lesions or contrast-enhancing lesions). MRI is also able to monitor aspects related to the neurodegenerative process typical of progressive MS, such as loss of brain volume, cortical lesion load, and the presence of chronically active lesions [19].

Thanks to its versatility and large numbers of imaging sequences, MRI is capable of deriving several imaging markers for MS.

Despite their importance, MRI markers can only partially explain the pathophysiological mechanisms underlying MS. Very often, their applicability to particular subgroups of patients is limited by the low specificity of standard neuroradiological protocols and by lack of semantic harmonization (scanner-dependent contrast scales) of the MRI images that makes them not comparable [37, 40].

Molecular biomarkers can excellently complement MRI, clinical, paraclinical, and instrumental data [45, 49]. This allows for earlier diagnosis and, therefore, treatment with drugs modifying the course of the disease in patients with clinically isolated syndrome (CIS). The advent of highly sensitive analytical platforms that allow reliable quantification of novel protein biomarkers present in low concentrations in peripheral blood samples has paved the way for future applications in clinical practice [1, 28]. At the moment, however, the use of these new biomarkers in clinical practice is limited by several factors, including low specificity, high individual variability also concerning physiological variables (age, body mass, renal function) and differences in measurement methods, high costs and poor accessibility [28]. Among the emerging biomarkers, there are also some CSF and serum cytokines related to neuroinflammation [4, 43].

Besides emerging serum biomarkers, additional clinical tools have been introduced to quantify different multidimensional aspects of MS, such as assessment of fatigue, cognition, and ambulation. The numerical quantification of these parameters, with relative objectification, could allow phenotyping MS in terms of disease activity or symptom-specific phenotypes [33, 36, 38, 39, 44].

In recent years, digital biomarkers have been introduced, consisting of measurements to collect objective data on biological (e.g., blood glucose, serum sodium), anatomical (e.g., mole size), or physiological (e.g., heart rate, blood pressure) parameters. The use of portable or implantable biosensors and their interpretable measures has allowed us to numerically evaluate important parameters such as, among others, cognitive functions and fatigue. Specific portable sensor-based measurement systems can be used in both clinical and domestic settings to carry out gait analysis (speed, acceleration, resistance, and quality of movement), posture, balance [39], as well as the fluidity of movement of the hands [36, 38]. In this way, it is possible to have a detailed vision of the daily life of patients and a quantitative and objective monitoring of some parameters correlated with the progression of MS.

Therefore, the "disease of a thousand faces" requires a complex, ubiquitous, and differentiated, as well as adaptive diagnosis, monitoring, and treatment strategy. This strategy should be personalized and adapted to the patient and the individualized course of the disease, and it should be continuously adapted to ensure that the disease progression is always the least fast.

Artificial intelligence (AI) promises to be a fundamental tool in the management of MS, especially in the field of automatic analysis

of the above multimodal and multidimensional data [6, 7]. AI could be well suited to predicting clinical disability, long-term benefits, and the safety of disease-modifying treatments through the multitude of tools it makes available. The outputs of the AI models, which can involve the analysis of distinct groups of data, can be reorganized together into a digital twin (DT). However, the adoption of AI is not without risks. As widely demonstrated, AI models may suffer from bias (i.e., unmotivated discrimination or favouritism of specific group of individuals [32]), privacy leaks of sensitive information [12], or lack of explainable predictions [29]. Hence, as also highlighted by the recently approved AI Act from the European Union [17], special attention must be devoted to ensuring the *ethical quality* of those AI systems (i.e., their fairness [11, 15, 16], privacy [13], and explainability [22, 23]).

In this paper, we propose the initial architecture of a DT specifically designed for MS (DTMS) that includes both the specificities of the patient and its relationship with the class (phenotype) of allowance in a combined DT. The architecture integrates and stores all the above-mentioned heterogeneous information into a so-called *semantically uniform database*. Those data are then fed as input to a federated AI model that is compliant with ethical guidelines to retrieve predictions about specific MS treatments for single patients or clusters of patients. In particular, the provided architecture and the envisaged DT will focus on improving both spatial and temporal accuracy to:

(1) Define new phenotypes (classes), or specific sub-phenotypes, in Figure 1;
(2) Provide an effective provisional follow-up strategy, based on combined digital markers, to ensure each patient is maintained in the minimal slope curve in Figure 1.

The rest of this paper is structured as follows: Section 2 presents some related work in the context of DT in the healthcare domain. Section 3 is devoted to describing the high-level architecture and the main components of our proposed approach. Finally, Section 4 discusses some future challenges and concludes the paper.

2 RELATED WORK

Digital twins in the healthcare domain are starting to be considered a well-established tool, but the advent of digital twins patients (DTPs) in personalized medicine comes with a new set of applications, benefits, and challenges [3]. A DTP is a virtual representation of the patient with whom to interact throughout life, which provides intelligence for diagnosis, evaluation, prediction, and treatment calibration. In its modern version, the DT has 5 dimensions: a physical entity, a digital counterpart, a connection that binds the two parts together, and the data and services [47].

A DT consists of numerous dynamic and multidimensional parameters [8, 25, 26]. Dynamic means that the data is both historically available and continuously updated and accumulated from that person's life, for example, related to the person's medical condition, diagnostics, living style and environment, tolerability of a drug, or degree of acceptance of therapy. The multidimensionality of the data comes from the many different sources generating them, e.g., a DTMS can be created to allow healthcare providers to manage diverse disease parameters, clinical and para-clinical outcomes, multi-omics, biomarkers, patient-related data, information on the patient's feelings and life plans, and medical procedures and protocols.

The potential, especially when it comes to precision medicine, is that DTs can be used to simulate individual therapies in advance and visualize potential therapeutic results in terms of progression, as treatments vary. The concept of DT seems to be particularly suitable for the approach to MS, in a DTMS, precisely because it is characterized by a heterogeneous course, complexity, and multidimensionality, a growing number of therapeutic options, and a consequent wealth of data [47].

This approach has been exemplified by studies such as the one by Steven Cen et al. on modeling disease-specific brain atrophy in individuals with multiple sclerosis [9]. Their research utilized a digital twin model to estimate the age of onset of brain atrophy by comparing longitudinal MRI data with a spline model derived from cross-sectional data of normal aging individuals [9].

Other works outline AI algorithms applied to analyze the data for diagnostic [2], personalized treatment recommendations, and identifying new MS phenotypes. Remote monitoring through wearable devices allows for continuous data collection and real-time monitoring, with continuous feedback loops in place to dynamically adjust and optimize treatment plans [48].

By unifying health monitoring, artificial intelligence models applied to MS [35], and simulations into fewer digital assets, DTPs enhance the interconnectivity of these modules. This holistic view allows for personalized health monitoring, identification of potential therapeutics, and overall improvement in patient care. DTPs can be instrumental in predicting adverse events, optimizing treatment plans, and facilitating large-scale clinical trials. Some of these studies are mentioned in a scoping review that summarizes patient digital twins, current applications and future directions while unfortunately highlighting that most of these projects are still in pre-clinical stages [14].

Addressing existing challenges [3] is essential for the broader adoption and effective implementation of digital twins in clinical practice. To advance, pragmatic and modular approaches must be established to develop the first clinical proof-of-concepts.

This work aims to go beyond the state of the art by combining both a patient-specific module and a phenotype-specific module that allows the definition of new digital markers for MS, while also predicting the temporal evolution of the disease. The DTMS aims to enhance the precision of MS management by providing real-time, individualized insights into disease progression and treatment efficacy.

3 DT CONCEPTS AND INITIAL ARCHITECTURE

Figure 2 shows the high-level architecture of the proposed approach. The system is designed as a layered architecture where data are constantly aggregated and processed from the bottom layer up to create digital twins (DT) for a single patient or a cluster of patients homogeneous for a given clinical demand of a specific type of MS (i.e., phenotype). The columns in the figure (C_1, C_2,...,C_k) represent the different clinical infrastructures and hospitals involved in the system.

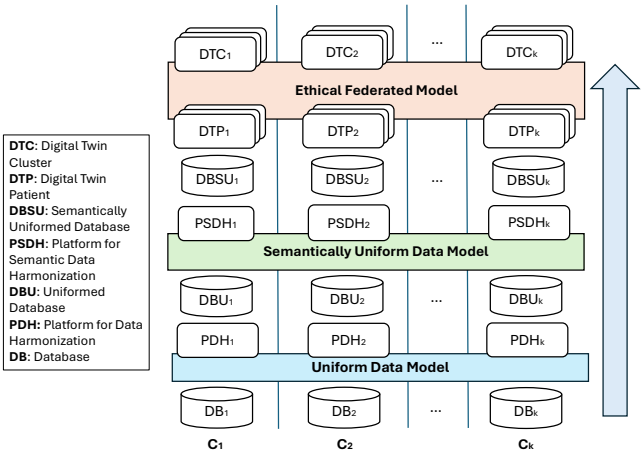

Figure 2: High level architecture

The base layer is identified by a series of large databases containing heterogeneous data coming from different clinical institutions (DB_1, DB_2, ..., DB_k in Figure 2). Given the heterogeneous nature of these data, they first have to be syntactically uniformed to be compliant with a uniform schema. Hence, the data are collected and preprocessed by a series of *Data Harmonization Platforms (PDH)* which model and uniform them following a *Uniform Data Model* schema. These uniformed data are then stored inside a series of *uniformed databases (DBU)*. However, these homogeneous data are still not ready to be fed as input to predict treatments for MS. Indeed, different systems for monitoring MS use different semantics to encode and represent information [42]; hence, the data has to be also semantically uniform to represent the same information in the same way. This operation is performed by a series of *Semantic Data Harmonization Platforms (PSDH)*, which collect the syntactically uniformed data from the *DBU* and perform a series of preprocessing operations to semantically uniform the data following a *Semantically Uniform Data Model*. This is particularly evident in the case of MRI. Each MRI manufacturer implements unique settings, leading to significant variability in the resulting images when the same study is conducted on the same patient using scanners from different producers. They are usually associated with the hardware and protocol used for capturing images. The properties of MRI, such as the strength of the scanner's magnetic field, the type of radiofrequency coil used, the characteristics of the gradients coil, the hardware, the image reconstruction method, and the parameters of the non-standardized acquisition process, can add technical variability that is also seen in the MRI-derived features [24, 27, 31].

Finally, these fully uniformed data are stored in a series of databases compliant with the semantically uniform schema (*DBSU* in Figure 2) and can be used as input for prediction tasks. In particular, the system will employ a federated learning model where the knowledge and information are shared among the machine learning (ML) models hosted in the different hospitals and clinical structures. The federated model will include *in-processing* fairness, explainability, and privacy models to ensure its reliability and trustworthiness, compliant with ethical guidelines for the adoption of ML systems in critical domains [17] (see Section 3.1). The predictions returned by the ML models, as well as the collected data, are finally used to generate the DT.

As said at the beginning of this section, the system provides two categories of DT: a DT tailored for a single patient (DTP in Figure 2) and a DT Aggregate tailored for a cluster of patients having the same clinical demand for a specific type of MS (DTC in Figure 2) [21]. In particular, the DTP will be generated by combining the information of a single patient with the predictions returned by the ML model. The information from all the patients exhibiting the same MS phenotype will be aggregated to generate the DTC. It is worth noticing how the two DTs will share information with each other. The data aggregated and collected from the DTC can be used in the DTP to improve its knowledge base and suggest actions to mitigate the progression of MS for all the patients of that cluster. At the same time, information from single patients can be used to detect new MS phenotypes.

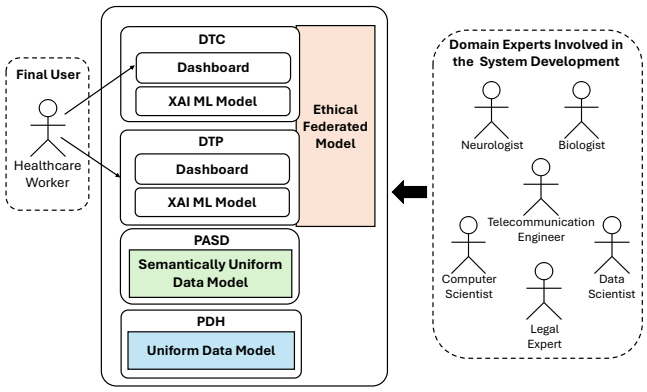

Figure 3: Detail of the DTs and main actors involved

Figure 3 reports a detailed view of the two types of DTs and the main actors involved in the system. As already highlighted in Figure 2, both DTs are based on the preprocessed and semantically uniform data provided by the PASD and stored in the DBSUs.[1] Both DTs comprise the same two macro-components: an XAI ML model (which is part of a larger federated model), which uses the data provided by the PASD to predict a therapeutic regimen for MS, and a dashboard to query the ML model and analyse the results. The final user of these DTs will be the healthcare worker. In particular, the healthcare worker will be able to interact with the DTC dashboard to analyse the predictions of the ML model concerning the cluster of patients or directly query the ML model of the DTP for a specific therapeutic regime. It is worth noticing how the two DTs will interact with each other. In particular, the DTC will collect the information from the underlying DTP in order to analyze common patterns for a cluster of patients. At the same time, if a patient moves from one cluster to another because of a change in the MS clinical scenario, then the DTP of that patient will be reconfigured using the information obtained from the new DTC. This communication between the two types of DTs will also allow the discovery of new clusters when the information from the DTP is not in line with the

[1] For figure clearness, the DBs are omitted in Figure 3.

existing ones. If several patient DTs show the same behavior, this opens a new challenge for the medical team because it should be the case that a new phenotype (patient cluster) has been identified. This feature can help improve the medical model in MS by potentially identifying new phenotypes (or patient clusters).

Finally, the right of Figure 3 details the main domain experts involved in the development of this system. As can be seen, different types of expertise are required, such as neurologists and biologists to provide domain knowledge about the treatment of MS, telecommunication engineers to provide knowledge about the technologies and data formats used to measure MS, computer and data scientists to provide knowledge about ML models and DTs development, and legal experts to provide legal knowledge to ensure the fairness and trustworthiness of the system.

3.1 Ethical requirements on DTMS

As already expressed in Section 1, given the sensitive application domain, special attention will be devoted to the ethical assurance of the DTMS. The ethical requirements we are addressing are privacy, fairness, and explainability [20].

Privacy is defined as *"the susceptibility of data or AI models to revealing private or sensitive information"* [10]. Concerning the privacy of data, the DBs shown in Figure 2 will all be hosted inside their relative institution, and no sensitive information will be exposed outside. Concerning the privacy of the underlying ML system, extensive research has been conducted to ensure the privacy of federated models [34]. Hence, during the development of the DTP, we will test several privacy-preserving models, like the one presented in [46], and select the one more suited for our use case. Concerning the DTC, privacy will be ensured by design since the data employed are aggregated.

Fairness is defined as *"the absence of any prejudice or favouritism toward an individual or group based on their inherent or acquired characteristics"* [32]. Fairness can be assessed either at an *individual* (i.e., all individuals must be treated equally despite any of their features) or *group* (all individuals must be treated equally despite their sensitive attributes like gender or ethnicity) level [11]. Since MS is related to some particular features of the patients and not all patients can be treated equally, we will focus on ensuring the fairness of this DTMS at a *group* level by applying in-processing fairness methods to ensure that the system does not apply discrimination towards any sensitive attribute (like gender, ethnicity, etc.).

Finally, in this context, we consider explainability as the ability to link any prediction of the ML model to features of the patient in the DT. This can be achieved in a federated learning context through a federation of explainable models [5].

4 DISCUSSION AND CONCLUSION

The chronic, heterogeneous, and multifocal nature of MS, also called the "disease of a thousand faces", requires a specific diagnostic strategy and differentiated, personalized, and adaptive monitoring and treatment. In this paper, we have proposed a novel DTMS designed to integrate diverse data sources, including MRI, clinical biomarkers, and digital health metrics, into a unified predictive model combining both a patient-specific module and a phenotype-specific module. The DTMS aims to enhance the precision of MS management by providing real-time, individualized insights into disease progression and treatment efficacy. The DTMS leverages explainable AI to offer reliable and personalized therapeutic recommendations, ultimately striving to delay disability and improve patient outcomes. This comprehensive digital framework represents a significant advancement in the application of AI and digital twins in the field of neurology, promising a more tailored and effective management strategy for MS and a predictive tool to anticipate the evolution of the disease. Future work will be the implementation of the proposed DTMS architecture, its integration with multimodal data and AI models, and, finally, its predictive validation with retrospective data from different specialistic neurological centers.

ACKNOWLEDGEMENTS

This research was funded by the European Union—NextGenerationEU under the Italian Ministry of University and Research (MUR) National Innovation Ecosystem grant ECS00000041—VITALITY—CUP E13C22001060006.

REFERENCES

[1] Ahmed Abdelhak, Matteo Foschi, Samir Abu-Rumeileh, John K Yue, Lucio D'Anna, Andre Huss, Patrick Oeckl, Albert C Ludolph, Jens Kuhle, Axel Petzold, et al. 2022. Blood GFAP as an emerging biomarker in brain and spinal cord disorders. *Nature Reviews Neurology* 18, 3 (2022), 158–172.

[2] H M Rehan Afzal, Suhuai Luo, Saadallah Ramadan, and Jeannette Lechner-Scott. 2022. The emerging role of artificial intelligence in multiple sclerosis imaging. *Mult. Scler.* 28, 6 (May 2022), 849–858.

[3] Patrizio Armeni, Irem Polat, Leonardo Maria De Rossi, Lorenzo Diaferia, Severino Meregalli, and Anna Gatti. 2022. Digital Twins in Healthcare: Is It the Beginning of a New Era of Evidence-Based Medicine? A Critical Review. *Journal of Personalized Medicine* 12, 8 (2022). https://doi.org/10.3390/jpm12081255

[4] Zhile Bai, Duanduan Chen, Luyao Wang, Yu Zhao, Tiantian Liu, Yun Yu, Tianyi Yan, and Yong Cheng. 2019. Cerebrospinal fluid and blood cytokines as biomarkers for multiple sclerosis: a systematic review and meta-analysis of 226 studies with 13,526 multiple sclerosis patients. *Frontiers in Neuroscience* 13 (2019), 1026.

[5] José Luis Corcuera Bárcena, Mattia Daole, Pietro Ducange, Francesco Marcelloni, Alessandro Renda, Fabrizio Ruffini, and Alessio Schiavo. 2022. Fed-XAI: Federated Learning of Explainable Artificial Intelligence Models.. In *XAI. it@ AI* IA*. 104–117.

[6] Andrea Bianchi and Antinisca Di Marco. 2024. Medical Knowledge Harmonization: A Graph-based, Entity-Selective Approach to Multi-source Diagnoses. In *The First International Conference on AI-Health (AIHealth 2024)*. Athens, Greece.

[7] Andrea Bianchi, Antinisca Di Marco, Francesca Marzi, Giovanni Stilo, Cristina Pellegrini, Stefano Masi, Alessandro Mengozzi, Agostino Virdis, Marco Salvatore Nobile, and Marta Simeoni. 2023. Trustworthy Machine Learning Predictions to Support Clinical Research and Decisions. In *2023 IEEE 36th International Symposium on Computer-Based Medical Systems (CBMS)*. IEEE, 231–234.

[8] Andrea Bianchi, Giordano d'Aloisio, Andrea D'Angelo, Antinisca Di Marco, Alessandro Di Matteo, Jessica Leone, Giulia Scoccia, Giovanni Stilo, and Luca Traini. 2022. DIORAMA: Digital twIn fOR sustAinable Territorial Management. In *In1st Italian Conference on Big Data and Data Science (ITADATA)*.

[9] Steven Cen, Mulugeta Gebregziabher, Saeed Moazami, Christina J. Azevedo, and Daniel Pelletier. 2023. Toward precision medicine using a "digital twin"approach: modeling the onset of disease-specific brain atrophy in individuals with multiple sclerosis. *Scientific Reports* 13, 1 (2023), 16279. https://doi.org/10.1038/s41598-023-43618-5

[10] Chris Clifton. 2009. Privacy Metrics. In *Encyclopedia of Database Systems*, LING LIU and M. TAMER ÖZSU (Eds.). Springer US, Boston, MA, 2137–2139. https://doi.org/10.1007/978-0-387-39940-9_272

[11] Giordano d'Aloisio, Claudio Di Sipio, Antinisca Di Marco, and Davide Di Ruscio. 2024. How fair are we? From conceptualization to automated assessment of fairness definitions. *arXiv preprint arXiv:2404.09919* (2024).

[12] Emiliano De Cristofaro. 2020. An overview of privacy in machine learning. *arXiv preprint arXiv:2005.08679* (2020).

[13] Davide Di Ruscio, Paola Inverardi, Patrizio Migliarini, and Phuong T Nguyen. 2024. Leveraging privacy profiles to empower users in the digital society. *Automated Software Engineering* 31, 1 (2024), 16.

[14] David Drummond and Apolline Gonsard. 2024. Patient digital twins: an introduction based on a scoping review. *medRxiv* (2024). https://doi.org/10.1101/2024.02.20.24303096

[15] Giordano d'Aloisio, Antinisca Di Marco, and Giovanni Stilo. 2023. Democratizing Quality-Based Machine Learning Development through Extended Feature Models. In *Fundamental Approaches to Software Engineering (Lecture Notes in Computer Science)*, Leen Lambers and Sebastián Uchitel (Eds.). Springer Nature Switzerland, Cham, 88–110. https://doi.org/10.1007/978-3-031-30826-0_5

[16] Giordano d'Aloisio, Andrea D'Angelo, Antinisca Di Marco, and Giovanni Stilo. 2023. Debiaser for Multiple Variables to enhance fairness in classification tasks. *Information Processing & Management* 60, 2 (2023), 103226.

[17] European Union. 2024. AI-Act. https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

[18] Massimo Filippi, Amit Bar-Or, Fredrik Piehl, et al. 2018. Multiple sclerosis. *Nature Reviews Disease Primers* 4 (2018), 43. https://doi.org/10.1038/s41572-018-0041-4

[19] Gavin Giovannoni, Veronica Popescu, Jens Wuerfel, Kerstin Hellwig, Ellen Iacobaeus, Michael B Jensen, José Manuel García-Domínguez, Livia Sousa, Nicola De Rossi, Raymond Hupperts, et al. 2022. Smouldering multiple sclerosis: the 'real MS'. *Therapeutic advances in neurological disorders* 15 (2022), 17562864211066751.

[20] Görkem Giray. 2021. A software engineering perspective on engineering machine learning systems: State of the art and challenges. *Journal of Systems and Software* (2021), 111031.

[21] Michael W Grieves. 2023. Digital twins: past, present, and future. In *The digital twin*. Springer, 97–121.

[22] Riccardo Guidotti, Anna Monreale, Fosca Giannotti, Dino Pedreschi, Salvatore Ruggieri, and Franco Turini. 2019. Factual and Counterfactual Explanations for Black Box Decision Making. *IEEE Intelligent Systems* 34, 6 (2019), 14–23. https://doi.org/10.1109/MIS.2019.2957223

[23] Riccardo Guidotti, Anna Monreale, Salvatore Ruggieri, Franco Turini, Fosca Giannotti, and Dino Pedreschi. 2018. A survey of methods for explaining black box models. *ACM computing surveys (CSUR)* 51, 5 (2018), 1–42.

[24] Xiao Han, Jorge Jovicich, David Salat, Andre van der Kouwe, Brian Quinn, Silvester Czanner, Evelina Busa, Jenni Pacheco, Marilyn Albert, Ronald Killiany, et al. 2006. Reliability of MRI-derived measurements of human cerebral cortical thickness: the effects of field strength, scanner upgrade and manufacturer. *Neuroimage* 32, 1 (2006), 180–194.

[25] Felicien Ihirwe, Davide Di Ruscio, Katia Di Blasio, Simone Gianfranceschi, and Alfonso Pierantonio. 2024. Supporting model-based safety analysis for safety-critical IoT systems. *Journal of Computer Languages* 78 (2024), 101243.

[26] Felicien Ihirwe, Davide Di Ruscio, Simone Gianfranceschi, and Alfonso Pierantonio. 2024. CHESSIoT: A model-driven approach for engineering multi-layered IoT systems. *Journal of Computer Languages* 78 (2024), 101254.

[27] Jorge Jovicich, Silvester Czanner, Douglas Greve, Elizabeth Haley, Andre van Der Kouwe, Randy Gollub, David Kennedy, Franz Schmitt, Gregory Brown, James MacFall, et al. 2006. Reliability in multi-site structural MRI studies: effects of gradient non-linearity correction on phantom and human data. *Neuroimage* 30, 2 (2006), 436–443.

[28] Michael Khalil, Charlotte E Teunissen, Sylvain Lehmann, Markus Otto, Fredrik Piehl, Tjalf Ziemssen, Stefan Bittner, Maria Pia Sormani, Thomas Gattringer, Samir Abu-Rumeileh, et al. 2024. Neurofilaments as biomarkers in neurological disorders—towards clinical application. *Nature Reviews Neurology* (2024), 1–19.

[29] Maximilian A. Kohl, Kevin Baum, Markus Langer, Daniel Oster, Timo Speith, and Dimitri Bohlender. 2019. Explainability as a Non-Functional Requirement. In *2019 IEEE 27th International Requirements Engineering Conference (RE)*. IEEE, Jeju Island, Korea (South), 363–368. https://doi.org/10.1109/RE.2019.00046

[30] Fred D Lublin, Stephen C Reingold, and National Multiple Sclerosis Society (USA) Advisory Committee on Clinical Trials of New Agents in Multiple Sclerosis*. 1996. Defining the clinical course of multiple sclerosis: results of an international survey. *Neurology* 46, 4 (1996), 907–911.

[31] Chiara Marzi, Marco Giannelli, Andrea Barucci, Carlo Tessa, Mario Mascalchi, and Stefano Diciotti. 2024. Efficacy of MRI data harmonization in the age of machine learning: a multicenter study across 36 datasets. *Scientific Data* 11, 1 (2024), 115.

[32] Ninareh Mehrabi, Fred Morstatter, Nripsuta Saxena, Kristina Lerman, and Aram Galstyan. 2021. A Survey on Bias and Fairness in Machine Learning. *Comput. Surveys* 54, 6 (July 2021), 1–35. https://doi.org/10.1145/3457607

[33] Sara Basso Moro, Silvia Bisconti, Makii Muthalib, Matteo Spezialetti, Simone Cutini, Marco Ferrari, Giuseppe Placidi, and Valentina Quaresima. 2014. A semi-immersive virtual reality incremental swing balance task activates prefrontal cortex: a functional near-infrared spectroscopy study. *Neuroimage* 85 (2014), 451–460.

[34] Viraaji Mothukuri, Reza M Parizi, Seyedamin Pouriyeh, Yan Huang, Ali Dehghantanha, and Gautam Srivastava. 2021. A survey on security and privacy of federated learning. *Future Generation Computer Systems* 115 (2021), 619–640.

[35] Yahya Naji, Mohamed Mahdaoui, Raymond Klevor, and Najib Kissani. 2023. Artificial intelligence and multiple sclerosis: Up-to-date review. *Cureus* 15, 9 (2023).

[36] Giuseppe Placidi. 2007. A smart virtual glove for the hand telerehabilitation. *Computers in Biology and Medicine* 37, 8 (2007), 1100–1107.

[37] Giuseppe Placidi. 2012. *MRI: essentials for innovative technologies*. CRC Press.

[38] Giuseppe Placidi, Danilo Avola, Luigi Cinque, Matteo Polsinelli, Eleni Theodoridou, and João Manuel RS Tavares. 2021. Data integration by two-sensors in a LEAP-based Virtual Glove for human-system interaction. *Multimedia Tools and Applications* 80 (2021), 18263–18277.

[39] Giuseppe Placidi, Danilo Avola, Marco Ferrari, Daniela Iacoviello, Andrea Petracca, Valentina Quaresima, and Matteo Spezialetti. 2014. A low-cost real time virtual system for postural stability assessment at home. *Computer methods and programs in biomedicine* 117, 2 (2014), 322–333.

[40] Giuseppe Placidi, Luigi Cinque, Matteo Polsinelli, et al. 2020. Guidelines for Effective Automatic Multiple Sclerosis Lesion Segmentation by Magnetic Resonance Imaging.. In *ICPRAM*. 570–577.

[41] Giuseppe Placidi, Mylene Sabatini, Massimo Gallucci, and Antonello Sotgiu. 2003. A calculation method for semi automatic follow up of multiple sclerosis by magnetic resonance eco planar perfusion imaging. In *The New Navigators: from Professionals to Patients*. IOS Press, 164–169.

[42] Àlex Rovira and Adelaida León. 2008. MR in the diagnosis and monitoring of multiple sclerosis: an overview. *European journal of radiology* 67, 3 (2008), 409–414.

[43] Foraysa Talaat, Sahar Abdelatty, Christine Ragaie, and Ahmed Dahshan. 2023. Chitinase-3-like 1-protein in CSF: a novel biomarker for progression in patients with multiple sclerosis. *Neurological Sciences* 44, 9 (2023), 3243–3252.

[44] Eleni Theodoridou, Luigi Cinque, Filippo Mignosi, Giuseppe Placidi, Matteo Polsinelli, João Manuel RS Tavares, and Matteo Spezialetti. 2022. Hand tracking and gesture recognition by multiple contactless sensors: A survey. *IEEE Transactions on Human-Machine Systems* 53, 1 (2022), 35–43.

[45] Alan J Thompson, Brenda L Banwell, Frederik Barkhof, William M Carroll, Timothy Coetzee, Giancarlo Comi, Jorge Correale, Franz Fazekas, Massimo Filippi, Mark S Freedman, et al. 2018. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *The Lancet Neurology* 17, 2 (2018), 162–173.

[46] Stacey Truex, Nathalie Baracaldo, Ali Anwar, Thomas Steinke, Heiko Ludwig, Rui Zhang, and Yi Zhou. 2019. A Hybrid Approach to Privacy-Preserving Federated Learning. In *Proceedings of the 12th ACM Workshop on Artificial Intelligence and Security* (London, United Kingdom) *(AISec'19)*. Association for Computing Machinery, New York, NY, USA, 1–11. https://doi.org/10.1145/3338501.3357370

[47] Isabel Voigt, Hernan Inojosa, Anja Dillenseger, Rocco Haase, Katja Akgün, and Tjalf Ziemssen. 2021. Digital twins for multiple sclerosis. *Frontiers in immunology* 12 (2021), 669811.

[48] Judith Wenk, Isabel Voigt, Hernan Inojosa, Hannes Schlieter, and Tjalf Ziemssen. 2024. Building digital patient pathways for the management and treatment of multiple sclerosis. *Frontiers in Immunology* 15 (2024). https://doi.org/10.3389/fimmu.2024.1356436

[49] Tjalf Ziemssen, Katja Akgün, and Wolfgang Brück. 2019. Molecular biomarkers in multiple sclerosis. *Journal of neuroinflammation* 16, 1 (2019), 272.