CoMoDID: Combining Explainable Artificial Intelligence and Conceptual Modeling for Data-Intensive Domains Management

Oscar Pastor[1,†], Alberto García S.[1,*,†], José Fabián Reyes Román[1,†], Diana Martínez Minguet[1,†], Mireia Costa[1,†], Ana León[1,†] and Ferrán Pla[1,†]

[1]*Valencian Research Institute for Artificial Intelligence (VRAIN). Universitat Politècnica de València, Camí de Vera S/N, Valencia, 46022, Spain*

Abstract

The large and heterogeneous data sets that characterize Data-Intensive Domains (DID) pose a challenge to developing data analysis and management approaches. A successful and efficient data-knowledge extraction from DID-based systems is determined by assembling and analyzing such data sets, but integrating their different sources is arduous work. Finding sound solutions for this problem has become a relevant research goal, as existing DID-based systems are not solving it convincingly. To solve this problem, a conceptual characterization of the data sets that constitute DID-based systems is essential. Using foundational ontologies and conceptual modeling provides an adequate strategy to face the complexity of this problem by clarifying the data structure that is to be analyzed and managed. In this project, we tackle this principle by defining a method grounded on a conceptual model to develop efficient DID-based systems and using a well-grounded combination of Explainable Artificial Intelligence (XAI) and Machine Learning (ML) techniques to perform data analytics. In addition, the characterization of a platform for implementing the method has been designed and developed. The project's chosen application domain is genomics, specifically in predicting critical diseases before symptoms manifest. Using XAI and ML with genomic information can contribute to the advancement of precision medicine, allowing the prediction of future diseases based on the available genomic data. The ML dimension covers the predictive knowledge (is a disease present in a patient?), while the XAI dimension deals with the explainable part (why does the patient have a disease?).

Keywords

Data-Intensive Domains, Conceptual Modeling, Explainable Artificial Intelligence, Precision Medicine

1. Introduction

Data has become an invaluable asset in today's society, and its production is unparalleled, continuously increasing. This presents significant challenges for modern software platforms that must store, analyze, and quickly provide access to data for numerous users. Consequently, various research fields related to data management and processing have undergone profound transformations [1]. One of the most current and relevant challenges in the software development context is dealing with DID-based systems, which require extensive and heterogeneous datasets [2] to create knowledge from data. Software developers must integrate complex, distributed, and heterogeneous datasets from increasingly diverse data-generating technologies (e.g., sensors, the Internet, genome sequencing machines, and other sophisticated devices) to develop effective and efficient methods and facilities for data analysis and management. Therefore, managing this massive amount of data to find the most critical and actionable

ER2025: Companion Proceedings of the 44th International Conference on Conceptual Modeling: Industrial Track, ER Forum, 8th SCME, Doctoral Consortium, Tutorials, Project Exhibitions, Posters and Demos, October 20-23, 2025, Poitiers, France

*Corresponding author.

†These authors contributed equally.

opastor@dsic.upv.es (O. Pastor); algarsi3@pros.upv.es (A. García S.); jreyes@pros.upv.es (J. F. Reyes Román); dmarmin@pros.upv.es (D. Martínez Minguet); miscossan@pros.upv.es (M. Costa); aleon@vrain.upv.es (A. León); fpla@dsic.upv.es (F. Pla)

0000-0002-1320-8471 (O. Pastor); 0000-0001-5910-4363 (A. García S.); 0000-0002-9598-1301 (J. F. Reyes Román); 0009-0002-3191-1969 (D. Martínez Minguet); 0000-0002-8614-0914 (M. Costa); 0000-0002-0877-7063 (A. León); 0000-0003-4822-8808 (F. Pla)

pieces of knowledge has become a significant challenge.

A fascinating example of DID-based systems is those that analyze the human genome [3]. Understanding the human genome is a significant scientific challenge, requiring the application of sound conceptual modeling techniques to manage such complex systems adequately. The continuous generation of genomic data from improved sequencing technologies [4, 5, 6] necessitates selecting the right data management strategy for software platforms. Developing software systems to deal with these DID is key for a proper genome analysis that would anticipate future illness in the human population [7].

To address these issues, this proposal is grounded on an interdisciplinary scientific policy, especially interested in combining two strong lines of research: conceptual modeling (CM) [8] and explainable artificial intelligence (XAI) [9]. To this aim, two main components are explored, designed, and developed: i) A method to deal with DIDs problem's management (the methodological perspective) correctly and efficiently, and ii) a "materialization" of the method in the form of a platform intended to assess the solution's value in a challenging and specially selected DID as the one related to the understanding of the human genome (the practical perspective).

In this scenario, applying a methodological framework based on XAI and CM to address DIDs concerns effectively becomes a relevant, promising strategy that forms the basis of the scientific approach used to achieve the project's primary goal. On the one hand, CM is recognized as crucial for developing data-oriented computer systems, ensuring an accurate representation of the application domain independently of the system that will be designed to address a real-world problem. This is especially relevant when we want to "understand data" in a DID context, which in our case applies to genomics. On the other hand, there is the application of XAI principles [9, 10], which describe a system in which humans can easily understand the results that an AI system provides, focusing primarily on understanding exactly "how" and "why" decisions are taken to reach results [11, 12]. For DID-based systems, where the proper representation of concepts becomes a crucial step, CM becomes the perfect partner for a practical XAI application [9] since by visualizing the relevant concepts, the structure of meaning people use to understand the domain is clearly represented. Our approach - both methodological (a method) and practical (a platform for the genomics domain)- is based on the group's expertise [13, 11], focusing on understanding data's true nature, employing CM techniques, and addressing challenges such as data volume and processing.

1.1. Details of the Project

The project combines XAI and CM for Data-Intensive Domains Management (CoMoDID). It is a four-year project (Sept. 2022 – Dec. 2025). Currently, the Research team is constituted by Óscar Pastor López, Juan Carlos Casamayor Ródenas, Tanja E. Vos, and Lluís-F. Hurtado, Encarna Segarra, Ferran Pla, Fernando García Granada, José F. Reyes Román (Postdoctoral Researcher), Alberto García Simón (Postdoctoral Researcher), Mireia Costa (Postdoctoral Researcher), and Diana Martínez Minguet (Predoctoral Researcher), in collaboration with the Genomics Team of the PROS research group. The Generalitat Valenciana supports the project through the CIPROM/2021/023 project.

2. Project Goals, Tangible Outputs & Expected Outcomes

The research proposed in this project focuses on the design of solutions for DID problems, as existing frameworks to build DID-based systems lack a solid conceptual modeling foundation and are built too frequently as ad hoc implementations. Both the method and the platform to materialize the solution, to show how it works for a selected DID, conform to the two primary objectives of this project:

- Definition of a general method (i.e., DELFOS) applicable to any DID for facing its analysis and design, which is based on a sound combination of conceptual modeling techniques and XAI technologies.
- Development of a technological platform that will instantiate and support the method in a particularly challenging and complex DID context: the genomic domain.

To address the methodological and practical components of the approach, we break these objectives into specific goals (G) with associated work packages (WPs), from which the tangible outputs obtained so far result.

2.1. Specific Goals for a General Method

The two specific goals to achieve our objective are:

G1 Ontological characterization of DIDs: This project achieved significant progress in the ontological characterization of Data-Intensive Domains. A complete and functional first version (metamodel) has been consolidated [14], which has been validated through two primary use cases: the modeling and specification of Clinical Management Systems (CMS) [15], and the analysis of genomic data associated with nuclear medicine and radiopharmacy [16].

Besides, we established a Framework for Ontology Conceptualization (F4OC) [17], which serves as a comprehensive set of best practices for ontological analysis of complex domains associated with DID platforms. This framework is supported by a meta-ontology focused on FAIR principles (Findable, Accessible, Interoperable, Reusable) [18].

The project also adapted foundational ontologies, such as UFO, to address CoMoDID's specific challenges through international collaborations with leading experts in ontological engineering [19] and conceptual modeling [20]. These efforts resulted in a modular proposal with precise ontological support for combining conceptual modeling and genomic data [21], facilitating proper integration of DIDs from diverse data sources. This approach has been evaluated with satisfactory results [22].

Most recent developments have focused on extending and consolidating the ontology framework, with particular emphasis on genomic data conceptual modeling [23] and DNA variant classification [24]. Relevant patterns for genomic data analysis have been identified and formalized [25], enabling the ontology to cover more specific and complex use cases within data-intensive domains. The ontological framework has been further presented to the scientific community and validated through the organization of the OntoCom workshop at JOWO 2024 [26] and culminated in a doctoral thesis [17] providing a comprehensive framework for ontological characterization with specific applications in cybersecurity domains, addressing crucial security and privacy aspects of DIDs.

G2 Integration of XAI techniques for data management and exploitation: This objective has been successfully completed through a comprehensive approach that progressed from initial comparative studies to the full development and validation of the DELFOS method.

Initial efforts focused on the genomic domain as a major DID, conducting statistical comparisons of different data sources with information associated with cancer and heart disease [27, 28, 29]. These initial studies established the basis for understanding data completeness and concordance challenges in genomic domains, leading to the development of XAI techniques to replicate clinical criteria for genomic data analysis and reduce manual expert activities.

Significant progress was achieved through international collaborations, resulting in the development of novel data management solutions. A Data Warehouse approach was implemented for genomic data extraction, integration, and storage, providing technological support for the DELFOS module, which is responsible for information storage and traceability throughout the analysis process. This development extended previous work on the Hermes platform for genomic data extraction, transformation, and integration [30]. The Conceptual Model of the Genome structure supported all of these scientific developments to ensure information quality through proper ontological support.

Throughout the CoMoDID project, the DELFOS method has undergone continuous improvement and validation, allowing its successful application to precision medicine scenarios [31]. The

method's effectiveness has been demonstrated through its application to early-onset Alzheimer's disease, validating its utility in identifying relevant variants in genomic information systems [32]. This validation represents a significant milestone in consolidating the methodological foundation and proving its practical applicability in concrete clinical scenarios.

Furthermore, the method's applicability has been extended to complex diseases, with particular focus on Polygenic Risk Scores and their translation to clinical practice through the DELFOS method [33]. This extension demonstrates the method's versatility and adaptability to different precision medicine contexts, completing the integration of XAI techniques for comprehensive data management and exploitation in DIDs.

2.2. Specific Goals for a Technological Platform

The instantiation of the method in the Genomics domain aims to validate that the method can be applied to complex DIDs, to provide a technological platform to collect, manage, and analyze the generated data in practical settings to improve the understanding of the human genome challenge, and ultimately to obtain relevant value by the extraction of knowledge from the data. To achieve these objectives, the following goals are defined:

G3 Definition of the interaction mechanisms for DID-based systems: This objective has been successfully accomplished through a comprehensive approach that evolved from initial interaction requirements analysis to the complete implementation of model-to-code transformations for DID-based systems.

Our initial work established the analysis of interaction requirements for DID-based systems and the elicitation of requirements for designing sustainable interfaces. This initial phase focused on developing comprehensive approaches for the conceptual modeling of genomic data and human-centered design principles for efficient management of smart genomic information, providing the theoretical and practical groundwork for defining interaction mechanisms.

We achieved significant progress by integrating the designed interaction model into a model-driven development method using an agile approach. This integration allowed us to reuse technologies developed by the research center to progress from interaction and business requirements to code generation [34]. This work incorporated an assistant for improving model comprehensibility [35]. This work also designed a method for integrating privacy and data protection requirements into user interfaces and the system in general [36].

The transformation mechanisms between abstract and concrete interaction models were successfully integrated into previously designed methods and technologies. The proposed design transforms business strategy and goal models (including sustainability considerations) through business process models, followed by information systems and interaction modeling that are processed to generate system code. This integrated approach has been validated in different iterations, starting from business strategy and goal levels [37], with particular attention to associating interaction models with requirements of other organizational actors through a proposed modeling method that includes a taxonomy of organizational actor types [38].

The completion of this objective was marked by significant advances in model-to-code transformations, as reported in the results of two doctoral theses providing complementary methodological frameworks: one focused on usability requirements capture [39] and another establishing a comprehensive model-driven software production method [40]. Applying Design Science principles [41] has provided the foundations for correct model-to-code transformation through an integral model-driven software production method that considers everything from strategy to code.

International collaboration has supported the development of these tools and methods through various undergraduate projects focused on business process model to class diagram transformations, class diagram to microservice code transformations, organizational modeling tools, information system tools, and business process model comprehensibility studies.

Advances in model-to-model transformations have complemented the work [42], demonstrating the viability of transforming models from one formal notation to another, and theoretical contributions on integrating fast and slow thinking in software development. Practical applicability has been demonstrated through systematic interpretation of variants [43, 25], completing the full spectrum of interaction mechanisms for DID-based systems.

G4 Development of a platform to support the DELFOS method: This objective has been successfully accomplished through the development of a platform composed of a suite of specialized tools for supporting the DELFOS method, addressing the complete workflow from data extraction to knowledge exploitation in genomic domains.

An initial work established comprehensive studies of medical literature repositories and analysis of information retrieval systems. Deep Learning models using Transformers were developed for multiclass and multilabel classification of medical radiology reports through Transfer Learning techniques using clinical and biomedical corpora.

Then, we systematically curated and analyzed biomedical scientific articles, identifying common clustering patterns and interaction relationships. Key clusters identified include chromosomal elements associated with diseases, diseases and their symptoms, chemical molecules and medications, and genetic variations. Critical interaction patterns were established between diseases and chemical molecules, genetic variations and diseases, and chromosomal elements and variations.

As a result of this work, we developed tools for automated extraction of gene-disease relationships from scientific literature [44], supported by specific methodologies for pattern detection in genomic information [45]. These tools were practically validated in familial cardiopathies, generating knowledge graphs connecting genes and associated diseases [46]. A semantics-based search engine was developed to address the fragmentation and heterogeneity of genomic information across diverse data sources, demonstrating significant improvements in search precision compared to traditional keyword-based methods, particularly validated in retina-macula diseases.

This suite of tools, integrated under the DELFOS platform [47], incorporates automated literature extraction capabilities, biological entity detection, and interaction analysis. Delfos includes a genomic database designed for storing and managing analyzed information, implemented following FAIR principles. Advanced exploitation modules were developed, including automatic genetic variation interpretation tools validated in hereditary retinal dystrophies, predictive systems for identifying genomic hotspots, and tools for genetic variant calling.

The platform architecture was implemented using modern microservices design, providing flexibility, scalability, and maintainability. Our tool has demonstrated versatility across different clinical domains, including integration with an information system for managing Neuroblastoma treatment data and its extension to complex diseases such as mental disorders [33, 48].

G5 DELFOS method and platform validation: This objective has been completed through a validation process that progressed from initial requirements identification to extensive real-world testing across multiple domains, demonstrating the robustness and adaptability of the DELFOS method and platform.

The validation process began with systematically identifying testing requirements for DID applications through collaborative meetings with domain experts. Initial validation was conducted through practical case studies of the DELFOS platform across different diseases [49]. Requirements analysis was extended to incorporate strategy and objectives analysis, ensuring compatibility with TestStar (Test*) technology [37].

Then, we designed a testing strategy for DID applications, addressing their unique characteristics. As a use case, it is focused on the genomic domain. The approach was based on automatic generation of multiple test cases using TestStar (Test*) technology, which was successfully integrated within a model-driven development method for agile software development contexts [37] and evaluated in real industrial settings [50].

The testing strategy underwent significant refinement, involving multiple working sessions with precision medicine professionals. This process resulted in diverse scientific outcomes, including research on polygenic risk score analysis [33], studies on how to apply DELFOS for dermatology, and oncology applications through collaborations with the Valencian Institute of Oncology (IVO). Finally, new testing techniques were developed to improve interface and report quality [51].

We demonstrated the method's applicability and effectiveness in a practical case study focused on early-onset Alzheimer's disease, followed by evaluations across multiple biological domains [32]. Successful validation was achieved in polygenic risk score optimization for complex diseases, including mental disorders such as autism and depression [33]. The method demonstrated its potential in managing, integrating, and analyzing clinical data and medical images using advanced AI and Machine Learning techniques in dermatology.

The validation scope extended beyond healthcare to demonstrate DELFOS's adaptability across different DID domains. Applications included software development for video games [52] and European projects measuring how music impacts people [53].

3. Relevance for ER

The proposed project is aligned with several research topics relevant to the conceptual modeling community. It is highly relevant to the issues of Ontological and cognitive foundations and Semantics in conceptual modeling since the incorporation of foundational ontologies and conceptual modeling in the project contributes to a solid theoretical foundation concerning DID-based systems, in combination with the project's focus on developing standardized approaches for data integration and analysis, which involves addressing semantic aspects. On the same line, the project is relevant to the topic of complexity management of large conceptual models, given that the project addresses the challenge of managing large and heterogeneous data sets in complex DID-based systems.

In another direction, the project aims to develop a method and platform to automate the development of DID-based systems, including data modeling. In this context, using Artificial Intelligence helps optimize and automate data analysis. However, in the Precision Medicine field, where the practical instantiation of the project is embedded, transparency and minimization of uncertainties are essential for the resulting decisions to be explainable. XAI satisfies these requirements, thus being suitable for data analysis counseling. The use of XAI and ML techniques involves knowledge representation and reasoning for accurate data analysis, which is directly related to logic-based knowledge representation and reasoning.

Overall, the proposed project's alignment with various research topics highlights its relevance and potential contributions to conceptual modeling, as well as knowledge representation and reasoning in the context of DIDs. It aims to address existing challenges and improve the efficiency and accuracy of DID-based systems, offering valuable insights for data analysts in diverse research fields based on conceptual modeling techniques and foundational ontologies.

4. Current Project Status

The project is approaching its completion, ending in the next four months. It has successfully achieved all primary objectives and progressed significantly, delivering comprehensive results across all work packages.

All primary objectives have been successfully completed: the ontological characterization of DIDs (G1) has been finalized with international validation through workshops and doctoral theses. The DELFOS method (G2) has been fully developed, validated, and applied to real clinical scenarios, including early-onset Alzheimer's disease and complex diseases. Interaction mechanisms for DID-based systems (G3) have been implemented with complete model-to-code transformations. A comprehensive set of tools and platforms (G4) has resulted in the fully functional DELFOS platform. Finally, extensive

validation (G5) has been conducted across multiple domains. We can conclude that the DELFOS platform has demonstrated its versatility through several successful implementations.

Current activities focus on generating final documentation, knowledge transfer, and consolidation of results for broader scientific dissemination.

Acknowledgments

We thank IIS-La Fe, INCLIVA, ISABIAL, Hospital Universitario Doctor Peset, and BIONOS for their massive contribution to this project and willingness to help. This work was supported by the Generalitat Valenciana through the CoMoDiD project (CIPROM/2021/023).

Declaration on Generative AI

During the preparation of this work, the author(s) used Grammarly in order to: Grammar and spelling check and reword. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the publication's content.

References

[1] A. Margara, G. Cugola, N. Felicioni, S. Cilloni, A Model and Survey of Distributed Data-Intensive Systems, 2023. doi:10.48550/arXiv.2203.10836, arXiv:2203.10836 [cs].

[2] A. Elizarov, B. Novikov, S. Stupnikov (Eds.), Data Analytics and Management in Data Intensive Domains: 21st International Conference, DAMDID/RCDL 2019, Kazan, Russia, October 15–18, 2019, Revised Selected Papers, volume 1223 of *Communications in Computer and Information Science*, Springer International Publishing, Cham, 2020. doi:10.1007/978-3-030-51913-1.

[3] M. Felderer, B. Russo, F. Auer, On Testing Data-Intensive Software Systems, in: S. Biffl, M. Eckhart, A. Lüder, E. Weippl (Eds.), Security and Quality in Cyber-Physical Systems Engineering: With Forewords by Robert M. Lee and Tom Gilb, Springer International Publishing, Cham, 2019, pp. 129–148. doi:10.1007/978-3-030-25312-7_6.

[4] K. R. Kumar, M. J. Cowley, R. L. Davis, Next-Generation Sequencing and Emerging Technologies, Seminars in Thrombosis and Hemostasis 45 (2019) 661–673. doi:10.1055/s-0039-1688446.

[5] D. J. Rigden, X. M. Fernández, The 27th annual Nucleic Acids Research database issue and molecular biology database collection, Nucleic Acids Research 48 (2020) D1–D8. doi:10.1093/nar/gkz1161.

[6] W. R. McCombie, J. D. McPherson, E. R. Mardis, Next-Generation Sequencing Technologies, Cold Spring Harbor Perspectives in Medicine 9 (2019) a036798. doi:10.1101/cshperspect.a036798.

[7] B. Louie, P. Mork, F. Martin-Sanchez, A. Halevy, P. Tarczy-Hornoch, Data integration and genomic medicine, Journal of Biomedical Informatics 40 (2007) 5–16. doi:10.1016/j.jbi.2006.02.007.

[8] A. Olivé, Conceptual Modeling of Information Systems, Springer, Berlin, Heidelberg, 2007. doi:10.1007/978-3-540-39390-0.

[9] O. P. López, A. L. Palacio, J. F. R. Román, J. C. Casamayor, Modeling Life: A Conceptual Schema-centric Approach to Understand the Genome, in: J. Cabot, C. Gómez, O. Pastor, M. R. Sancho, E. Teniente (Eds.), Conceptual Modeling Perspectives, Springer International Publishing, Cham, 2017, pp. 25–40. doi:10.1007/978-3-319-67271-7_3.

[10] A. Barredo Arrieta, N. Díaz-Rodríguez, J. Del Ser, A. Bennetot, S. Tabik, A. Barbado, S. Garcia, S. Gil-Lopez, D. Molina, R. Benjamins, R. Chatila, F. Herrera, Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI, Information Fusion 58 (2020) 82–115. doi:10.1016/j.inffus.2019.12.012.

[11] S. Spreeuwenberg, Choose for AI and for Explainability, in: C. Debruyne, H. Panetto, W. Guédria, P. Bollen, I. Ciuciu, G. Karabatis, R. Meersman (Eds.), On the Move to Meaningful Internet Systems:

OTM 2019 Workshops, Springer International Publishing, Cham, 2020, pp. 3–8. doi:10.1007/978-3-030-40907-4_1.

[12] J. Qin, M. Žumer, X. Wang, W. Fan, Conceptual Models and Ontological Schemas for Semantically Sustainable Digital Libraries, in: Proceedings of the ACM/IEEE Joint Conference on Digital Libraries in 2020, JCDL '20, Association for Computing Machinery, New York, NY, USA, 2020, pp. 441–442. doi:10.1145/3383583.3398545.

[13] L. Kalinichenko, A. Volnova, E. Gordov, K. Nadezhda, D. Kovaleva, O. Malkov, I. Okladnikov, N. Podkolodnyy, A. Pozanenko, N. Ponomareva, S. Stupnikov, A. Fazliev, Data access challenges for data intensive research in Russia, Informatics and Applications 10 (2016) 2–22. doi:10.14357/19922264160101.

[14] M. Costa, A. G. S., A. Leon, A. Bernasconi, O. Pastor, A Reference Meta-model to Understand DNA Variant Interpretation Guidelines, in: Conceptual Modeling: 42nd International Conference, ER 2023, Lisbon, Portugal, November 6–9, 2023, Proceedings, Springer-Verlag, Berlin, Heidelberg, 2023, pp. 375–393. doi:10.1007/978-3-031-47262-6_20.

[15] M. Costa, A. García S., A. León, O. Pastor, Comprehensive Representation of Variation Interpretation Data via Conceptual Modeling, in: T. P. Sales, J. Araújo, J. Borbinha, G. Guizzardi (Eds.), Advances in Conceptual Modeling, Springer Nature Switzerland, Cham, 2023, pp. 25–34. doi:10.1007/978-3-031-47112-4_3.

[16] S. Arevshatyan, J. F. Reyes Román, E. Caballero Calabuig, M. C. Plancha, A. Abella, P. Abreu, O. Pastor, Integrating Nuclear Medicine and Radiopharmacy Data: A Conceptual Model for Precision Medicine and Enhanced Patient Care, in: Advances in Conceptual Modeling: ER 2023 Workshops, CMLS, CMOMM4FAIR, EmpER, JUSMOD, OntoCom, QUAMES, and SmartFood, Lisbon, Portugal, November 6–9, 2023, Proceedings, Springer-Verlag, Berlin, Heidelberg, 2023, pp. 15–24. doi:10.1007/978-3-031-47112-4_2.

[17] B. Franco Martins Souza, A Framework for Conceptual Characterization of Ontologies and its Application in the Cybersecurity Domain (2024). URL: https://riunet.upv.es/handle/10251/204584, publisher: Universitat Politècnica de València.

[18] B. F. Martins, J. F. R. Román, O. Pastor, M. Hadad, Improving Conceptual Domain Characterization in Ontology Networks, in: S. Nurcan, A. L. Opdahl, H. Mouratidis, A. Tsohou (Eds.), Research Challenges in Information Science: Information Science and the Connected World, Springer Nature Switzerland, Cham, 2023, pp. 187–202. doi:10.1007/978-3-031-33080-3_12.

[19] A. García Simón, Understanding the Code of Life: Holistic Conceptual Modeling of the Genome (2023). URL: https://riunet.upv.es/handle/10251/191432, publisher: Universitat Politècnica de València.

[20] A. Bernasconi, G. Guizzardi, O. Pastor, V. C. Storey, Semantic interoperability: ontological unpacking of a viral conceptual model, BMC Bioinformatics 23 (2022) 491. doi:10.1186/s12859-022-05022-0.

[21] A. Bernasconi, A. García S., S. Ceri, O. Pastor, PoliViews: A comprehensive and modular approach to the conceptual modeling of genomic data, Data & Knowledge Engineering 147 (2023) 102201. doi:10.1016/j.datak.2023.102201.

[22] A. García S., A. Bernasconi, G. Guizzardi, O. Pastor, V. C. Storey, I. Panach, Assessing the value of ontologically unpacking a conceptual model for human genomics, Information Systems 118 (2023) 102242. doi:10.1016/j.is.2023.102242.

[23] J. I. Panach, O. Pastor, S. W. Liddle, V. C. Storey, H. C. Mayr, B. Thalheim, Evaluating a Framework of Conceptual Modelling Research, in: Advances in Conceptual Modeling: ER 2024 Workshops, AISA, CMLS, EmpER, QUAMES, JUSMOD, LLM4Modeling, Pittsburgh, PA, USA, October 28–31, 2024, Proceedings, Springer-Verlag, Berlin, Heidelberg, 2024, pp. 209–218. doi:10.1007/978-3-031-75599-6_16.

[24] M. Costa, A. García S., A. León, A. Bernasconi, O. Pastor, VarClaMM: A reference meta-model to understand DNA variant classification, Data & Knowledge Engineering 154 (2024) 102370. doi:10.1016/j.datak.2024.102370.

[25] M. Costa, O. Pastor, A Conceptual Approach to Using Relevant Patterns in Genomic Data Analysis,

ER Forum, Special Topics, Posters and Demos Co-located with ER 2024 (2024).
[26] I. J. da Silva Oliveira, P. P. F. Barcelos, R. F. Calhau, C. M. Fonseca, et al. (Eds.), Proceedings of the Joint Ontology Workshops (JOWO) - Episode X: The Tukker Zomer of Ontology, and satellite events co-located with the 14th International Conference on Formal Ontology in Information Systems (FOIS 2024), Enschede, The Netherlands, July 15-19, 2024, volume 3882 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2024. URL: https://ceur-ws.org/Vol-3882.
[27] A. García S., M. Costa, A. Leon, O. Pastor, The challenge of managing the evolution of genomics data over time: a conceptual model-based approach, BMC Bioinformatics 23 (2022) 472. doi:10.1186/s12859-022-04944-z.
[28] M. Costa, A. García S., O. Pastor, A Comparative Analysis of the Completeness and Concordance of Data Sources with Cancer-Associated Information, in: R. Guizzardi, B. Neumayr (Eds.), Advances in Conceptual Modeling, Springer International Publishing, Cham, 2022, pp. 35–44. doi:10.1007/978-3-031-22036-4_4.
[29] M. Costa, A. García S., O. Pastor, Conceptual Modeling-Based Cardiopathies Data Management, in: R. Guizzardi, B. Neumayr (Eds.), Advances in Conceptual Modeling, Springer International Publishing, Cham, 2022, pp. 15–24. doi:10.1007/978-3-031-22036-4_2.
[30] S. G. Fernandes, hermes: Uma platform para a extraçaõ, trasformaçaõ e integraçaõ de dados genomicos, univesidade Do Minho (2022).
[31] P. Ferreira, Delfos: Designing and implementing a distributed architecture to store and analyse genomics data., univesidade Do Minho (2022).
[32] M. Costa, A. León, O. Pastor, Improving the identification of relevant variants in genome information systems: A methodological approach with a case study on early onset Alzheimer's disease, Data & Knowledge Engineering 151 (2024) 102284. doi:10.1016/j.datak.2024.102284.
[33] D. Martínez-Minguet, O. Pastor, Conceptual Modeling for Polygenic Risk Score Research: Improving Domain Understanding and Clinical Utility, in: Advances in Conceptual Modeling: ER 2024 Workshops, AISA, CMLS, EmpER, QUAMES, JUSMOD, LLM4Modeling, Pittsburgh, PA, USA, October 28–31, 2024, Proceedings, Springer-Verlag, Berlin, Heidelberg, 2024, pp. 159–168. doi:10.1007/978-3-031-75599-6_12.
[34] R. Noel, R. Velasquez, J. I. Panach, O. Pastor, Agile development of Systems from REquirements to Code (SREC), RPE@CAiSE'23: Research Projects Exhibition at the International Conference on Advanced Information Systems Engi- neering, June 12–16, 2023, Zaragoza, Spain (2023).
[35] R. Velasquez, C. Negri-Ribalta, R. Noel, O. Pastor, Exploring Understandability in Socio-technical Models for Data Protection Analysis: Results from a Focus Group, in: T. P. Sales, J. Araújo, J. Borbinha, G. Guizzardi (Eds.), Advances in Conceptual Modeling, volume 14319, Springer Nature Switzerland, Cham, 2023, pp. 263–273. doi:10.1007/978-3-031-47112-4_25, series Title: Lecture Notes in Computer Science.
[36] C. Negri-Ribalta, R. Noel, O. Pastor, C. Salinesi, An Empirical Study on Socio-technical Modeling for Interdisciplinary Privacy Requirements, in: M. Sellami, M.-E. Vidal, B. Van Dongen, W. Gaaloul, H. Panetto (Eds.), Cooperative Information Systems, volume 14353, Springer Nature Switzerland, Cham, 2024, pp. 137–156. doi:10.1007/978-3-031-46846-9_8, series Title: Lecture Notes in Computer Science.
[37] R. Noel, J. I. Panach, O. Pastor, Including business strategy in model-driven methods: an experiment, Requirements Engineering 28 (2023) 411–440. doi:10.1007/s00766-023-00400-3.
[38] R. Noel, J. I. P. Navarrete, O. Pastor, Using Team Topologies in Model-Driven Strategic Alignment, in: Congresso Ibero-Americano em Engenharia de Software (CIbSE), SBC, 2023, pp. 168–175. doi:10.5753/cibse.2023.24701, iSSN: 0000-0000.
[39] O. Ayala, Y. Isela, Marco para la captura de requisitos de usabilidad en entornos de MDD (2024). URL: https://riunet.upv.es/handle/10251/202630, publisher: Universitat Politècnica de València.
[40] N. López, R. Alejandro, Design and Implementation of a Model-Driven Software Production Method: from Strategy to Code (2024). URL: https://riunet.upv.es/handle/10251/203527, publisher: Universitat Politècnica de València.
[41] O. Pastor, M. A. Segooa, J. I. Panach, Teaching Design Science as a Method for Effective Research

Development, in: D. Mendez, P. Avgeriou, M. Kalinowski, N. B. Ali (Eds.), Handbook on Teaching Empirical Software Engineering, Springer Nature Switzerland, Cham, 2024, pp. 479–499. doi:10.1007/978-3-031-71769-7_17.

[42] A. León, M. Y. Santos, A. García, J. C. Casamayor, O. Pastor, Model-to-Model Transformation, Business & Information Systems Engineering 66 (2024) 85–110. doi:10.1007/s12599-023-00824-9.

[43] M. Costa, SISO: A Conceptual Model-based Method for Variant Interpretation Systematization, ER2024: Companion Proceedings of the 43rd International Conference on Conceptual Modeling: ER Forum, Special Topics, Posters and Demos (2024).

[44] L. Contreras-Ochando, P. Marco Garcia, A. León, L.-F. Hurtado, F. Pla, E. Segarra, Enhancing Precision Medicine: An Automatic Pipeline Approach for Exploring Genetic Variant-Disease Literature, in: T. P. Sales, J. Araújo, J. Borbinha, G. Guizzardi (Eds.), Advances in Conceptual Modeling, Springer Nature Switzerland, Cham, 2023, pp. 35–43. doi:10.1007/978-3-031-47112-4_4.

[45] M. Costa, O. Pastor, A Conceptual Approach to Using Relevant Patterns in Genomic Data Analysis, ER2024: Companion Proceedings of the 43rd International Conference on Conceptual Modeling: ER Forum, Special Topics, Posters and Demos, October 28-31, 2024, Pittsburgh, Pennsylvania, USA (2024).

[46] A. García S., M. Costa, A. Perez, O. Pastor, CardioGraph: a platform to study variations associated with familiar cardiopathies, BMC Medical Informatics and Decision Making 23 (2024) 303. doi:10.1186/s12911-024-02700-2.

[47] A. L. Palacio, A. G. S, J. F. R. Roman, M. Costa, O. Pastor, The Delfos Platform: A Conceptual Model-Based Solution for the Enhancement of Precision Medicine, IEEE/ACM transactions on computational biology and bioinformatics 21 (2024) 1242–1253. doi:10.1109/TCBB.2024.3377928.

[48] D. Martínez-Minguet, A. G. Simón, M. Costa, O. Pastor, Facilitating Genetic Risk Assessment of Mental Disorders, 11th ACM Celebration of Women in Computing: womENcourage 2024, Madrid, Spain (2024).

[49] J. F. Reyes Román, A. León Palacio, A. García Simón, R. C. Beyrouti, O. Pastor, Integration of clinical and genomic data to enhance precision medicine: a case of study applied to the retina-macula, Softw. Syst. Model. 22 (2022) 159–174. doi:10.1007/s10270-022-01039-4.

[50] A. Bons, B. Marín, P. Aho, T. E. J. Vos, Scripted and scriptless GUI testing for web applications: An industrial case, Information and Software Technology 158 (2023) 107172. doi:10.1016/j.infsof.2023.107172.

[51] F. Pastor Ricós, B. Marín, I. S. Prasetya, T. E. Vos, J. Davidson, K. Hovorka, An Industrial Experience Leveraging the iv4XR Framework for BDD Testing of a 3D Sandbox Game: 18th International Conference on Research Challenges in Information Science, RCIS 2024, Research Challenges in Information Science - 18th International Conference, RCIS 2024, Proceedings (2024) 393–409. doi:10.1007/978-3-031-59465-6_24, publisher: Springer.

[52] D. Blasco Latorre, Towards Search-based Game Software Engineering (2024). URL: https://riunet.upv.es/handle/10251/203655, publisher: Universitat Politècnica de València.

[53] European Commission. Directorate General for Education, Youth, Sport and Culture., KEA European Affairs., Panteia., Feasibility study for the establishment of a European Music Observatory: final report., Publications Office, LU, 2020. URL: https://data.europa.eu/doi/10.2766/9691.