CSSDH: A Unified Ontology for Social Determinants of Health and Seamless Continuity of Care Data Exchange

Subhashis Das*, Debashis Naskar, Sara Rodríguez González

BISITE Research Group, Department of Computer Sciences and Automation, University of Salamanca, Salamanca (Spain)

* Corresponding author: subhashis@usal.es

Received 3 April 2025 | Accepted 24 July 2026 | Early Access 28 August 2026

Abstract

The growth of digital platforms has accelerated the shift toward technology-driven, home-based healthcare solutions, empowering individuals to monitor their health and share data with healthcare professionals when necessary. However, developing an effective care plan management system involves more than simply analyzing hospital records and Electronic Health Records (EHRs). It requires addressing individual patient needs and considering social determinants of health, such as living conditions and the exchange of healthcare information across different care settings. This complex healthcare ecosystem faces challenges like schema diversity (across EHRs, personal health records, etc.) and varying medical terminologies (e.g., ICD, SNOMED-CT) used in ancillary healthcare services. Achieving interoperability between diverse systems and applications is essential. The European Interoperability Framework (EIF) highlights the importance of ensuring patient-centric access and control over healthcare data. In this paper, we introduce an integrated ontological model—the Common Semantic Data Model for Social Determinants of Health (CSSDH)—which combines ISO/DIS 13940:2025 ContSys with the WHO framework for Social Determinants of Health. CSSDH is designed to enhance interoperability across the continuity of care networks.

Keywords

Continuity of Care, EHR, ISO, Knowledge Graph, Ontology, OWL, Social Determinants of Health.

DOI: 10.9781/ijimai.2026.2236
eLocator: e2236

I. Introduction

The rapid expansion of digital platforms and services is playing a transformative role in narrowing the gap between organizational objectives and the diverse needs of citizens—especially in the domains of health and social care. This digital shift highlights the growing significance of remote monitoring technologies and home-based care solutions, which are driving an increasing demand for accessible and digitized healthcare services. Yet, while the adoption of Electronic Health Records (EHRs) has become widespread, reliance on these systems alone is not enough to ensure comprehensive, person-centered care. A deeper understanding of individual patient needs, along with the broader social context in which care is delivered, remains critical.

Key factors that influence the successful delivery of digital health services include procurement flexibility, socioeconomic disparities, levels of digital literacy, and informed consent regarding the sharing and use of personal health information. As healthcare delivery continues to move across various settings—from hospitals to homes to community centers—the ability to securely access and exchange health data requires a robust and interoperable healthcare management infrastructure. This includes advanced access controls, comprehensive data modeling, and strict privacy safeguards to ensure that sensitive patient information is both protected and usable where it is needed most.

According to the World Health Organization's (WHO) findings [1], poorly designed and inefficient EHR systems frequently contribute to a significant knowledge gap among policymakers, healthcare professionals, and researchers. These gaps hinder efforts to collect, analyze, and act on the data needed to improve population health outcomes. Without effective design and integration, critical insights remain out of reach, limiting the impact of healthcare interventions. WHO emphasizes the importance of leveraging diverse data sources—including primary care records, acute and home care data, national registries, and vital statistics systems—to provide a more comprehensive understanding of health trends and needs. Such data sources are essential for adapting information systems to support areas like rehabilitation services, where multi-faceted and evolving data needs are common.

The WHO's 2024 report, *Operational Framework for Monitoring Social Determinants of Health Equity (SDHE)* [2], further underscores the urgent need to transform public health data systems. The inequities observed during the COVID-19 pandemic in exposure rates, illness

Please cite this article as:
S. Das, D Naskar, S. Rodríguez González. CSSDH: A Unified Ontology for Social Determinants of Health and Seamless Continuity of Care Data Exchange, *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 10, no. 1, e2236, 2026, http://doi.org/10.9781/ijimai.2026.2236

severity, and mortality have laid bare the systemic gaps in existing data collection and monitoring frameworks. These gaps prevent health systems from effectively addressing social inequities and highlight the necessity of equity-oriented data systems. By focusing on the Social Determinants of Health (SDH), public health systems can better monitor factors that significantly influence health outcomes.

It is widely recognized that SDH contribute between 30% and 55% of overall health outcomes, underscoring their profound impact on individuals and communities. As healthcare continues to move toward more holistic, continuous models of patient care, capturing and integrating SDH data into health information systems is no longer optional—it is essential. If EHR systems are not equipped to record and track social determinants such as housing stability, education, employment status, or access to nutritious food, healthcare providers will lack a complete picture of a patient's well-being.

Recent initiatives, such as the U.S. Department of Health and Human Services' *Healthy People 2030* [3] strategy, emphasize the importance of systematically capturing SDH data to inform care planning, resource allocation, and policy decisions. By embedding SDH data into EHRs and other health information platforms, healthcare organizations can better address health disparities, improve patient outcomes, and create more equitable healthcare systems. Finally, the evolution of digital health solutions presents significant opportunities to enhance care delivery, but realizing their full potential requires a comprehensive approach. This involves not only improving the technological infrastructure but also prioritizing an understanding of patients' social and economic contexts. Doing so will ensure that healthcare systems are responsive, equitable, and capable of delivering high-quality care to all individuals, regardless of their circumstances. This work addresses the following research question: *How can social determinants of health be formally represented within continuity of care models in a way that supports semantic interoperability, automated reasoning, and practical deployment in real-world healthcare systems?* To answer this question, we design a unified ontology (CSSDH), implement it in a semantic knowledge graph, and evaluate its ability to support reasoning, competency question answering, and SDH-aware care analysis.

In this paper, we introduce an integrated, ontology-based information framework known as the *Common Semantic Data Model for Social Determinants of Health* (CSSDH). This comprehensive model is developed by harmonizing and extending two established standards: ISO 13940, which defines the System of Concepts to Support Continuity of Care (ContSys), and the World Health Organization's framework on Social Determinants of Health (WHO SDH). By unifying these conceptual models, the CSSDH facilitates semantic interoperability and enhances the standardized exchange of data related to social determinants of health across different healthcare systems. Ultimately, this model aims to support continuity of care by enabling consistent, accurate, and meaningful sharing of information between diverse healthcare providers, systems, and stakeholders. From an AI perspective, CSSDH leverages description logic–based reasoning to enable automated knowledge discovery, inference, and explainable decision support over heterogeneous healthcare data. Apart from the model, we have also presented a prototype mobile app, which is able to collect SDH data in a more structured way. The paper is structured as follows: related work is presented in Section 2, the overall methodology is described in Section 3, implementation of the model is presented in Section 4, and results and evaluation, along with a discussion on future work, are covered in Section 5.

This paper is an extended version of the work originally published at the 25th International Conference on Intelligent Data Engineering and Automated Learning (IDEAL) conference 2024 titled "CSSDH: An Ontology for Social Determinants of Health to Operational Continuity of Care Data Interoperability" [4]. The extended version elaborated in detail the technical system architecture to generate a personalized knowledge graph.

II. Related Work and Shortcoming

Current literature reveals a limited number of initiatives dedicated to formalizing models that enable semantic interoperability for continuity of care, particularly those that integrate social care dimensions into EHR systems. While there is growing recognition of the importance of incorporating social determinants of health into healthcare data models, practical and formalized approaches remain limited.

One noteworthy example is the *Health at Home* project [5], which proposed a conceptually integrated framework aimed at embedding social care considerations throughout the patient journey. Similarly, Cantor and Thorpe [6] have emphasized the critical need to integrate SDH data into EHR systems from a population health management perspective, identifying key challenges such as data standardization, interoperability, and the ethical handling of sensitive social data.

From the perspective of ontology engineering, the field is still in its infancy. Till date, only two ontologies directly addressing SDH that are available in the BioPortal [7] repository: the Social Determinants of Health Ontology (SOHO) [8] and the Social Prescribing Ontology [9]. However, both are primarily structured as simple term lists in Web Ontology Language (OWL) format and lack the formal precision and comprehensive structure needed to serve as full-fledged information models suitable for data integration across complex healthcare systems. Their limited expressiveness constrains their utility in supporting interoperability and continuity of care use cases. Furthermore, while the SDoHo ontology [10] offers a broader scope, it includes numerous classes that are not specific to SDH—such as demographic details and general patient attributes like age—resulting in conceptual dilution and reduced precision for SDH-focused applications.

Although there have been few localized efforts to establish universal screening mechanisms for SDH data collection alongside EHR systems [11], these initiatives have not yet yielded comprehensive information models or publicly available tools that facilitate standardized data sharing and interoperability. To the best of our knowledge, no integrated and formal ontological models currently exist that fully incorporate SDH into healthcare information models in a manner consistent with established interoperability frameworks.

In our prior work [12], we addressed this gap by developing a formal ontological model grounded in the ISO/DIS 13940:2025 ContSys standard [13] and aligned with the top-level ontology Descriptive ontology for linguistic and cognitive engineering (DOLCE) [14]. Our model provides a structured, semantically rich representation of continuity of care concepts, facilitating interoperability within and across healthcare systems. To the best of our knowledge, no integrated and formal ontological models currently exist.

To better position the proposed CSSDH ontology with respect to existing work, Table I presents a comparative analysis against representative state-of-the-art ontologies addressing SDH and related healthcare domains. While existing approaches primarily focus on SDH terminology, social prescribing workflows, or isolated demographic factors, they generally lack formal integration with continuity of care models and offer limited support for automated reasoning and deployment. In contrast, CSSDH combines standardized SDH representations with a formally grounded continuity of care ontology, enabling complex SDH-aware queries, description-logic-based reasoning, and practical knowledge graph implementation. CSSDH focuses on semantic interoperability and reasoning rather than predictive modeling; however, it provides a structured knowledge foundation upon which AI-driven analytics and decision support systems can be built.

TABLE I. Comparison of CSSDH With Representative State-of-the-Art SDH and Continuity-of-Care Ontologies

Feature	SOHO	SP Ont.	SDoHo	CSSDH
Primary focus	SDH terms	Social Rx	SDH + demo	SDH + CoC
Ontological grounding	Low	Low	Partial	Strong
WHO SDH aligned	Partial	Partial	Yes	Yes
ContSys compliant	No	No	No	Yes
ICD-11 / Gravity	No	No	Partial	Yes
OWL formalization	Low	Low	Med.	High
Automated reasoning	Limited	Limited	Partial	Yes
Complex SDH queries	Limited	Limited	Partial	Yes
SDH first-class	Yes	Yes	Yes	Yes
KG implementation	No	No	No	Yes
CQ-based validation	No	No	Partial	Yes
Privacy-aware SDH	No	No	No	Yes
Deployment shown	No	Partial	No	Yes
AI capabilities	Limited	Limited	Partial	Strong

In this paper, we extend our previously proposed model [15] to incorporate concepts explicitly related to social determinants of health. This extension aims to bridge the gap between clinical care and social care by enabling the seamless integration of SDH data into continuity of care frameworks. Our enhanced model supports a more holistic understanding of patient contexts, ultimately contributing to improved care coordination, health outcomes, and equitable healthcare delivery. Simultaneously, we aim to capture SDH concepts in a more structured manner and integrate them with the existing EHR system using a unique patient identifier.

In contrast to existing SDH ontologies that primarily act as terminological resources, and continuity-of-care models that focus mainly on clinical workflows, CSSDH uniquely combines (i) formal ontological grounding in top-level ontology DOLCE, (ii) alignment with ISO 13940-System of concepts to support continuity of care, (iii) standardized SDH vocabularies, and (iv) deployable reasoning infrastructure. To the best of our knowledge, no existing work integrates all these aspects into a single, operational semantic framework. We also have developed a prototype mobile application, which is described in detail in the implementation section IV.

III. Methodology

This paper presents the development of a formal ontological schema—called the Common Semantic Data Model for Social Determinants of Health (CSSDH)—designed to represent social factors impacting patients' continuous clinical outcomes. The following sections detail the methodology used to create CSSDH.

A. Modeling Objectives and Design Principles

The design of the proposed ontology is guided by a set of modeling objectives aimed at supporting the integration of social determinants of health within continuity of care scenarios while preserving semantic precision and interoperability. A primary objective is to incorporate SDH concepts without altering or weakening the semantics of the underlying ContSys model, ensuring compatibility with established continuity of care standards. To address privacy and data availability constraints commonly associated with SDH information, the ontology adopts a privacy-aware modeling strategy in which certain SDH attributes are represented using Boolean or categorical abstractions rather than fine-grained personal details. In line with the principle of reuse over reinvention, the ontology aligns its core concepts with existing standards and vocabularies, including FHIR, ICD-11, WHO SDH frameworks, and previously established SDH ontologies, thereby maximizing interoperability and facilitating adoption. Finally, all modeling decisions are made with reasoning compatibility in mind, ensuring that the resulting ontology supports efficient description-logic reasoning and can be effectively deployed within a GraphDB [16] environment for automated inference, complex querying, and personalized knowledge graph construction.

B. Modelling Concepts, Object Properties and Data Properties

Full formal ontology for continuity of care comprise of around 171 classes. Out of that main classes needed to capture our scenario are rather small. They are:

- *healthcare actor*: organization or person participating in healthcare. The involvement of the healthcare actor will be either direct (for example, the actual provision of care), or indirect (for example, at organizational level).
- *subject of care*: healthcare actor with a person role; who seeks to receive, is receiving, or has received healthcare. Synonym: subject of healthcare; service user; patient; client.
- *healthcare profession*: having a healthcare professional entitlement recognized in a given jurisdiction. The healthcare professional entitlement entitles a healthcare professional to provide healthcare independent of a role in a healthcare organization.
- *observation*: observations are a central element in healthcare, used to support diagnosis, monitor progress, determine baselines and patterns, and even capture demographic characteristics. We reuse HL7 FHIR observation [17] Resource in our ontology which provide details data structure for capturing patient daily observation.

The *Observation* class is responsible for capturing various measurements and making fundamental assertions about a subject, which could be a patient, a medical device, or another entity under observation. In the context of patient data, the *subject* of the observation is represented by the *Patient* class, while the *performer* of the observation is an entity within the *Healthcare Professional* class.

One key aspect of the *Observation* class is the restriction applied to certain data properties to ensure data consistency and accuracy. Specifically, the *person.gender* property is subject to a datatype restriction, requiring data providers to select from predefined gender values: *"Male," "Female," or "Transsexual."* This restriction is mandatory, as gender information is considered essential for service providers in delivering appropriate care.

Furthermore, the system enforces strict validation rules:

- The *gender field cannot be left blank*; it is a required attribute.
- The *reasoning mechanism* within the system is designed to detect and flag any incorrect or invalid gender values entered into the system, ensuring compliance with predefined constraints.

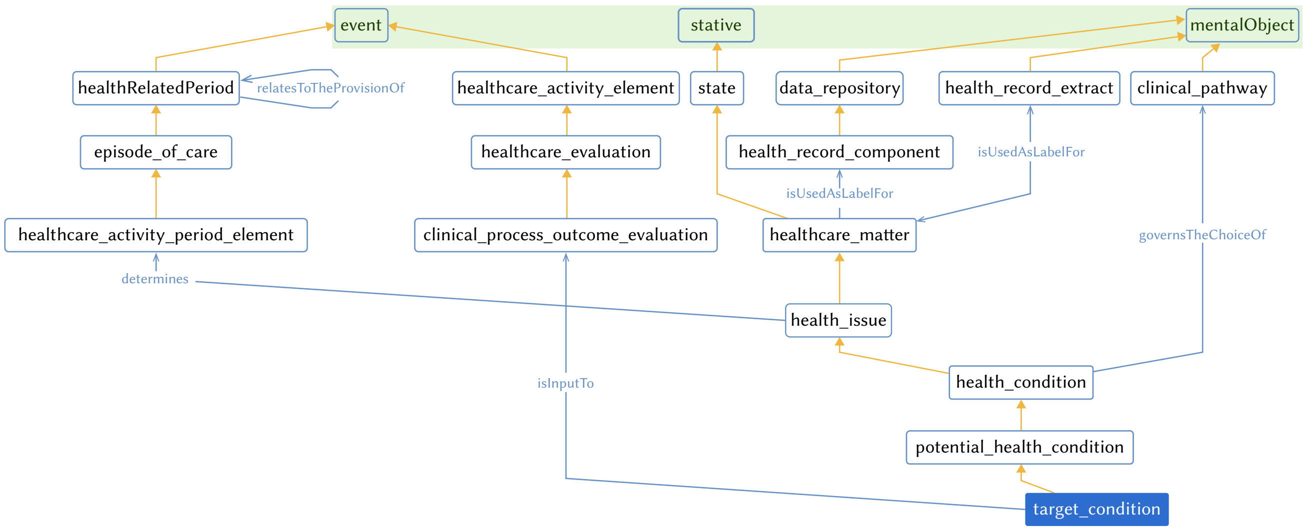


Fig. 1. target condition Ñ health Condition alignment with DOLCE Top-level (DOLCE classes are in green). The yellow line corresponds to IS_A relation, and the blue line is for the object-object relation.

The following RDF Turtle syntax example illustrates how this *Observation* class is structured and how the *gender property* is encapsulated within the data model.

```
[basicstyle=]
http://purl.org/net/for-coc#Observation
CoC:Observation rdf:type owl:Class;
rdfs:subClassOf [rdf:type owl:Restriction;
owl:onProperty
observation-definitions:Observation.performer;
owl:someValuesFrom EWS:healthcare_professional],
[rdf:type owl:Restriction;
owl:onProperty observation-definitions:
Observation.subject;
owl:someValuesFrom CoC:Patient];
oboInOwl:hasDbXref
https://www.hl7.org/fhir/observation-definitions.html
#Observation^^xsd:anyURI;
rdfs:comment Measurements and simple assertions
made about a patient,device or
other subject.@en.
```

person.gender(property)

```
https://www.hl7.org/fhir/person-definitions.html#
Person.gender
person-definitions:Person.gender
rdf:type owl:DatatypeProperty;
rdfs:range [rdf:type rdfs:Datatype;
owl:unionOf ([rdf:type rdfs:Datatype;
owl:oneOf [rdf:type rdf:List;
rdf:first Female^^xsd:string;
rdf:rest rdf:nil]]
[rdf:type rdfs:Datatype;
owl:oneOf [rdf:type rdf:List;
rdf:first Male^^xsd:string;
rdf:rest rdf:nil]]
[rdf:type rdfs:Datatype;
owl:oneOf [rdf:type rdf:List;
rdf:first Transsexual^^xsd:string;
rdf:rest rdf:nil]])];
rdfs:comment gender might not match the biological
sex as determined by genetics.
```

Fig. 1 illustrates a partial view of our ContSys ontology class structure, visualized using the *WebProtégé* platform. The upper section of the figure (highlighted in green) displays DOLCE foundational classes such as *mentalObject*, *stative*, and *event*, which provide the ontological grounding. The remaining classes are domain-specific, derived from the ContSys standard, and tailored to represent healthcare-specific processes and entities. In this study, we utilized four primary resources to support the collection and organization of concepts related to social determinants of health. These resources include Chapter 24 of the International Classification of Diseases, 11^{th} Revision (ICD-11), the Social and Healthcare Ontology (SOHO) [8], the Social Prescribing Ontology [9], and the SDH value set provided by the Gravity Project[1]. Drawing extensively from the SOHO ontology, we reused much of its existing terminology as a foundation for our model. To ensure a logically consistent and well-structured hierarchy, we organized these concepts in accordance with the OntoClean methodology [18].

For instance, the concept *Lay Off From Job,* which falls under the *Economic Stability* category, has been modeled as a data property linked to the *subject of care* class. This allows for the precise collection of relevant data regarding an individual's employment status. In addition to reusing established terms, we have extended our model by incorporating several new properties to enhance its comprehensiveness and granularity. Examples of these additions include *Medical Services Not Available at Home, Problems Associated with Exposure to Radiation,* and *Problems Associated with Exposure to Tobacco Smoke.* These terms, borrowed from ICD-11, were integrated to support the detailed documentation and analysis of various environmental and healthcare-related challenges that can impact health outcomes. By leveraging these diverse resources and ontological frameworks, our approach aims to create a more standardized and interoperable structure for capturing social determinants of health, ultimately facilitating better data collection and more informed decision-making in healthcare settings.

C. Ontology Alignment

Top-level ontologies offer a domain-independent framework for conceptualization by defining generalized categories, relations, and axioms. These categories—such as *event*, *mental object*, *quality*, and others—serve as foundational building blocks to standardize ontology development. By providing rigorous ontological commitments and guidelines, top-level ontologies promote consistency and

[1] https://thegravityproject.net/

interoperability across different knowledge domains. Consequently, constructing a *Continuity of Care* ontology within the structure of a top-level ontology enhances its potential for semantic interoperability with other established ontological models in the healthcare domain. *Descriptive Ontology for Linguistic and Cognitive Engineering* (DOLCE) [14], is an upper-level ontology designed to capture the fundamental categories of reality as they are perceived and conceptualized by humans. DOLCE is developed by Nicola Guarino and his team at Institute of Cognitive Sciences and Technologies (ISTC), CNR, Italy as part of the WonderWeb project [19], DOLCE emphasizes cognitive and linguistic aspects rather than strictly mirroring the physical world. In our work, we adopt the DOLCE as a middle-out approach, striking a balance between formalization and complexity. In 2023, DOLCE was adopted as an *ISO standard—ISO/IEC 21838-3:2023 Information technology — Top-level ontologies (TLO), Part 3* [20] —enhancing its status as a model for international acceptance.

This choice facilitates the creation of an ontology that is both theoretically sound and practically applicable. Despite the significant advantages that top-level ontologies offer, their implementation and alignment present non-trivial challenges and often demand the expertise of seasoned ontologists. Several initiatives, such as the European Union project *Advancing Clinico-Genomic Trials* (ACGT) [21] and other healthcare interoperability projects, have underscored the importance and benefits of aligning domain ontologies with top-level frameworks. In our *Continuity of Care* ontology, key DOLCE classes—such as *stative*, *mental object*, *event*, *natural person*, and *physical object* form the upper tiers of our class hierarchy. This hierarchical structure helps to organize and aggregate semantically similar classes, promoting coherence and facilitating interoperability with other ontologies that adopt comparable design patterns. Moreover, the need for rigorous ontological scrutiny of models like ContSys (the ISO 13940 standard for continuity of care) has been emphasized by the European Federation for Medical Informatics (EFMI) in their position paper [22]. However, progress in advancing such formal ontological alignment has been limited over the past decade.

Our formal ontology model addresses these longstanding issues by collaborating closely with the ISO 13940 community. Through this collaborative effort, we aim to refine and extend the continuity of care model with enhanced semantic clarity and interoperability. A fragment of our formal OWL ontology, developed in the Protégé environment, is illustrated in Fig. 1, showcasing the integration of DOLCE top-level classes within the structure of our continuity of care model. Below we have describe how CSSDH ontology is align with the DOLCE top-level concepts.

- event: something that happens at a given place and time. Within continuity of care domain concepts such as episode of care, hospital admission and adverse event are all specialization (i.e. subclass/subordinate class) of Event type.
- stative: expressing existence or a state rather than an action. Health condition which is an aspect of a health state is categorized under the top-level concept stative.
- mental Object: the sum or range of what has been perceived, discovered, or learned. Within continuity of care concept like care plan, clinical guideline, health record etc are classified under mental object.
- role: aka 'functional role'. A concept that classifies (in particular, it is 'played by') endurants, as used in some description. Roles are the descriptive counterpart of endurants, and, as endurants participate in perdurants, they usually have courses as modal targets [23]. Role is fundamental concept which interconnected various components within care deliver. Different persons play different role in a care setting such as doctor, patient, healthcare organization, next of kin or even subject of care proxy who make decision on behalf of a patient all are categorized under role.

IV. Implementation

The CSSDH ontology is developed using Protégé [24], an open-source ontology editing and management tool maintained by the Stanford Center for Biomedical Informatics Research. Protégé provides a robust environment for building and visualizing ontologies, allowing for the precise definition of concepts, relationships, and data properties necessary for modeling complex domains such as social determinants of health.

The primary objective of the CSSDH ontology is to *capture and represent patients' social determinants of health data within electronic health record systems*, while maintaining strict adherence to *patient privacy and data security principles*. By integrating SDH data into EHRs, healthcare professionals can gain a more comprehensive understanding of the non-clinical factors that influence patient health outcomes. At the same time, the model is carefully designed to minimize risks associated with handling sensitive personal information.

A key feature of the CSSDH ontology is its ability to answer *Competency Questions (CQs)* [25], which are predefined queries designed to validate the ontology's usefulness and scope. Examples of such competency questions include:

- `CQ1: Retrieve all patients who live in a low-income area.`
- `CQ2:Provide a list of patients who have been laid off from their jobs and whoalso live in overcrowded housing conditions.`
- `CQ3:Provide a list of patients who have Anxiety from community violence as well as food insecurity.`
- `CQ4: GeoSpatial Distribution of Prescribed Quantity for All Patients.`

These queries demonstrate the ontology's capability to support complex reasoning and facilitate data-driven decision-making in clinical and public health contexts.

The CSSDH ontology currently comprises *171 classes*, representing key entities such as *Subject of Care* (i.e., the patient), *Care Professional*, *Observation*, *Hospital Appointment*, and *Referral*. In addition, it includes *141 object properties*, which define the relationships between these classes. To capture specific attributes and data points, we have implemented a total of *210 data properties*, of which *171 are directly related to SDH factors*. These data properties cover a wide range of domains, including economic stability, education, social and community context, healthcare access, and neighborhood and built environment.

Given the sensitive nature of SDH data, we made a deliberate design decision to *represent most SDH information as Boolean values (i.e., true/false)*. This binary approach simplifies data representation and enhances privacy protection by limiting the granularity of sensitive personal information. This strategy aligns with established privacy and security guidelines, recognizing that detailed SDH data, if not carefully handled, could pose significant privacy risks [12].

As illustrated in Fig. 2, the *Subject of Care* class is shown within the class hierarchy panel of Protégé , along with its associated property restrictions. The right-hand panel provides a detailed view of SDH-related data properties linked to this class. These include categories such as *Social and Community Context* and *Neighborhood and Built Environment*, which represent key determinants influencing patient health.

This phase of the project centers on facilitating close collaboration among key stakeholders, including advanced nurse practitioners, healthcare professionals, local IT teams, and data scientists. Together, this multidisciplinary team works to co-design and develop a

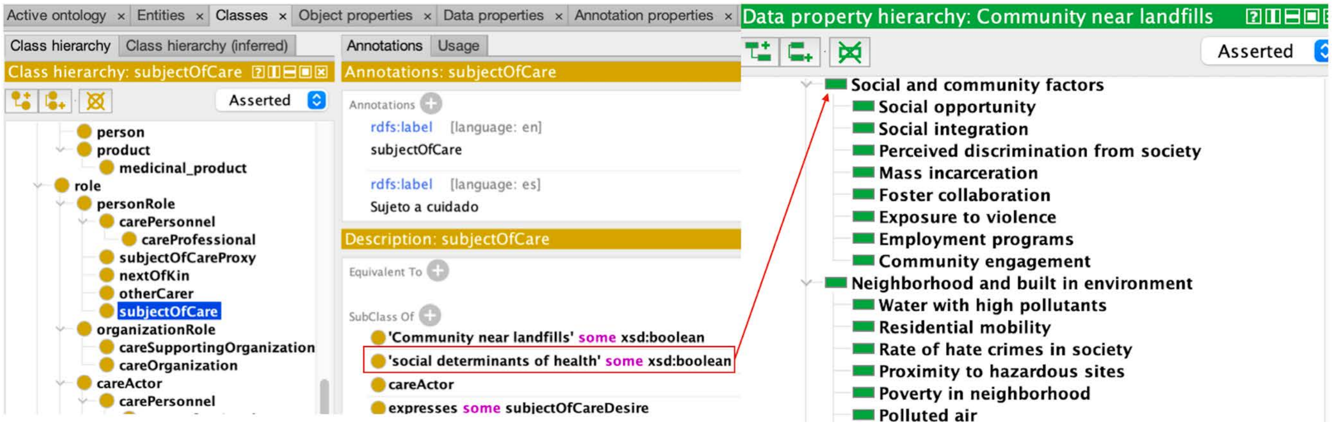


Fig. 2. Subject of Care SDH.

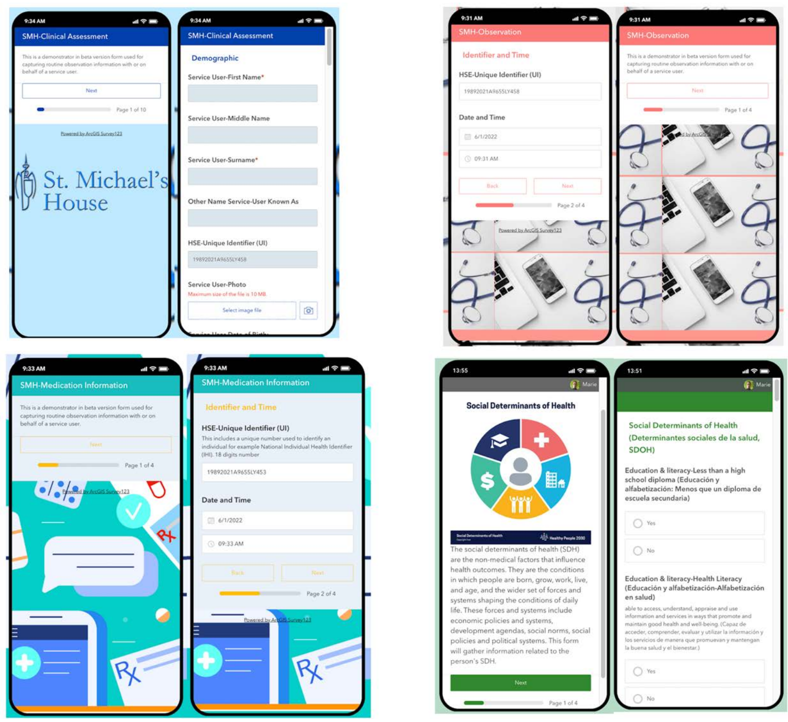


Fig. 3. Prototype Mobile App for Data Entry and Collection.

functional demonstrator aimed at streamlining healthcare data collection processes. This collaborative effort focuses on creating intuitive, electronic data collection forms that can systematically capture essential healthcare information in a structured, standardized format (see Fig. 3). The prototype mobile application is designed to capture structured patient data, including information related to social determinants of health. The underlying data model is aligned with a top-level ontology DOLCE, enabling clinicians to visualize all relevant aspects of a patient's clinical and social context in an integrated manner. This unified view supports informed decision-making during the prescribing process by allowing physicians to consider both medical and non-clinical factors that may influence treatment outcomes. The objective is to move beyond traditional, paper-based workflows—referred to as a "Model of Use"—and transition toward a more sophisticated "Model of Meaning." Informal qualitative feedback was obtained from participating healthcare professionals during the design of the data collection forms and SDH concepts. Experts confirmed that the structured representation of SDH factors aligns with real clinical workflows and improves the interpretability and reuse of social data compared to free-text documentation. In this new model, a cloud-based demonstrator enables seamless machine-to-machine communication, supporting interoperability and data integration across systems [26].

A critical aspect of this process involves deconstructing complex information into discrete, clearly defined data fields, rather than allowing data to remain embedded in free-text boxes. For instance, address details that might typically be recorded in a single text field are instead broken out into standardized components—such as Street Name, Apartment or Unit Number, City, and Postal Code—following best practices commonly seen in platforms like Amazon [2].

During the design of these electronic forms, special attention is given to determining the hierarchy and relevance of data elements. Stakeholders collaboratively define which data points are considered *core* (essential across all contexts), which are *common* (frequently used but not universally required), and which are *context-specific* (relevant to specific organizations or care settings only) [27]. This structured approach ensures the data collected is both meaningful and actionable.

[2] https://www.amazon.es/

Once consensus is reached on the data elements and their structure, the information is implemented into the demonstrator platform. This prototype—illustrated in Fig. 3 showcases a mobile application specifically developed for use in Irish-based residential care settings. The demonstrator not only facilitates accurate data entry at the point of care but also lays the groundwork for scalable, interoperable digital health solutions. Later on, the App instance capturing social determinants of health is extended and enhanced with a Spanish lemma to better address the local needs of the Castilla y León region in Spain.

By leveraging ontology-driven modeling and adhering to privacy-centric principles, the CSSDH ontology provides a structured, interoperable framework for capturing and analyzing SDH data in clinical systems. This model supports healthcare providers in identifying social risk factors and designing more personalized and holistic care interventions.

To populate CSSDH model with data, we utilized the free version of Ontotext GraphDB [16] Version 10.6 to support our system's semantic data management. GraphDB employs a hierarchical Role-Based Access Control (RBAC) model, which enables administrators to assign specific roles to users based on their level of authorization during server setup. The primary roles available include:

- ADMIN: Grants full control over all operations without restrictions.
- USER: Allows saving SPARQL queries, creating graph visualizations, and storing user-specific settings.
- MONITORING: Enables real-time monitoring of system operations, including executing queries, tracking updates, aborting queries or updates, and monitoring resource usage.
- REPO MANAGER: Provides the ability to create, modify, and delete repositories, as well as read and write access to all repositories within the system.

Additionally, GraphDB supports integration with the Lightweight Directory Access Protocol (LDAP), enhancing security by enabling centralized authentication and user management.

This robust access control mechanism, combined with LDAP integration, ensures the security of the healthcare system by mitigating unauthorized third-party access. Furthermore, the CSSDH model can be deployed on commercial cloud service providers, such as *Amazon Neptune*, where security and compliance measures are managed by the cloud provider—such as Amazon Web Services (AWS). AWS offers built-in security features, including data encryption, identity and access management (IAM), and network security controls, ensuring compliance with industry standards.

The practical safety, reliability, and regulatory compliance of CSSDH ultimately depend on the implementing healthcare system provider. In our deployment, this responsibility is managed by *Davra*, an Ireland-based technology startup specializing in large-scale CSSDH implementations. Davra adheres to multiple regulatory compliance frameworks to ensure data privacy, security, and reliability, including: FedRAMP (Federal Risk and Authorization Management Program), *HIPAA* (Health Insurance Portability and Accountability Act) Compliance, *ISO 27001* (International Standard for Information Security Management), *HITRUST* (Health Information Trust Alliance) Certification and *NIST 8259 CSF2014* (National Institute of Standards and Technology Cybersecurity Framework for Cloud Software).

By leveraging Davra's expertise and compliance with these stringent security standards, our deployment of CSSDH ensures the highest levels of data protection and regulatory alignment, making it a secure and scalable solution for healthcare applications.

V. Results and Discussion

The evaluation presented in this section aims to assess whether the proposed CSSDH ontology effectively answers the stated research question by enabling reasoning, complex querying, and SDH-driven insights in continuity of care scenarios.

A. Validation Objectives

The validation of CSSDH focuses on assessing whether the ontology:

1. supports complex SDH-related queries,
2. enables automated reasoning and inference, and
3. is usable and meaningful for healthcare professionals in real-world settings.

The Web Ontology Language (OWL) provides not only the ability to model relationships between different classes and individuals but also includes comprehensive *schema information*, such as *class hierarchies, property characteristics, and logical constraints*. This additional layer of formalized knowledge enables more advanced capabilities beyond simple data representation. Specifically, OWL supports *automated reasoning*, allowing users to infer new knowledge, validate data consistency, and uncover implicit relationships that are not explicitly stated in the knowledge base.

By leveraging *description logic-based reasoning*, OWL ontologies can be subjected to *logical inference processes.* These processes systematically apply schema-based rules and axioms to the existing data, generating new facts or conclusions that extend the original knowledge base. Inference mechanisms are crucial for tasks such as *classification* (automatically organizing individuals into appropriate classes), *consistency checking* (detecting logical contradictions in the data), and *relationship discovery* (revealing hidden connections). The expressiveness of our CSSDH ontology model is ALCHQ(D) as per the description logic (DL) scale [28]. We have not exploited the full power of DL-full as supported by the OWL-2 language; rather, we used simple rules to make our model compatible with the GraphDB rule engine and offer quick query execution times.

B. Reasoning and Inference Validation

In the context of our case study, the reasoning capabilities of OWL proved particularly beneficial. By applying automated reasoning over the ontology, we were able to *derive new insights* about the modeled domain and *detect potential inconsistencies* in the continuity of care data. This not only enhanced the quality and reliability of the knowledge base but also facilitated more informed decision-making.

An example of automated inference is demonstrated in Fig. 4, where the HermiT reasoner [29] is used to classify entities, infer implicit relationships, and validate the ontology. The left panel depicts the inferred class hierarchy as visualized in the Protégé ontology editor. The main top-level classes in the hierarchy include *event, nonPhysicalObject, physicalObject, role,* and *stative*. The class *subjectOfCare* is modeled as a specialization of *personRole* and is positioned within the *role* subclass hierarchy. The right panel shows the Description Logic (DL) query interface, where a logical query is executed to retrieve all subjects of care who have been laid off from their jobs and are also living in overcrowded housing conditions. The query results identify two patients—John Murphy and Mary Smith—who satisfy both conditions, thereby illustrating the ontology's ability to support automated reasoning and complex SDH-based queries.

The reasoning process automatically populated new assertions based on the logical definitions and constraints specified in the OWL schema, showcasing how semantic technologies can enhance data interoperability and integrity within healthcare information

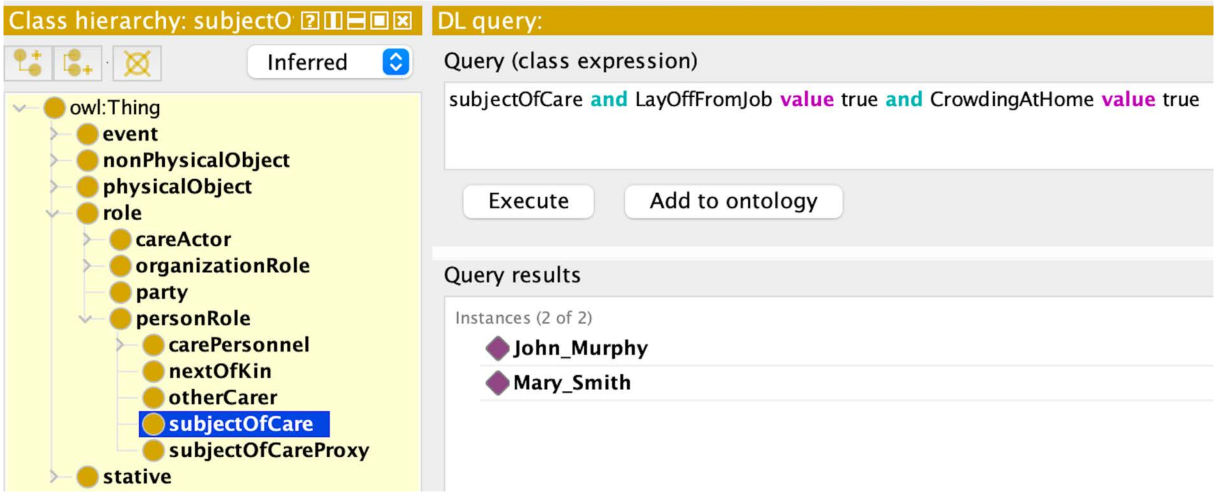


Fig. 4. Description Logic (DL) query execution with HermiT reasoner.

systems. While large-scale performance benchmarking is outside the scope of this paper, the ontology was deliberately designed using the ALCHQ(D) description logic profile to remain compatible with GraphDB's rule-based reasoner, ensuring efficient query execution and reasoning performance in practical deployments. Automated reasoning enables the discovery of implicit patient groupings and care-relevant conditions that are not explicitly asserted in the data, such as identifying socially vulnerable patient cohorts based on combinations of SDH factors.

C. Functional Validation Via Competency Questions

Competency Questions (CQs) are used as a functional validation mechanism to assess whether CSSDH can represent and retrieve SDH-informed clinical knowledge, as commonly recommended in ontology engineering. *CQ2 combines economic stability + housing conditions. CQ3 combines mental health + community violence + food insecurity.* To address *CQ2*, we associated the data properties *LayOffFromJob* and *CrowdingAtHome* with the class *subjectOfCare* and set their data types as `Boolean`. A snapshot of the `SPARQL` query structure is provided.

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX  coc: <http://purl.org/net/for-coc#>
PREFIX fhir: <http://hl7.org/fhir/>
SELECT ?subjectOfcare
 ?forename ?Lay_off_from_job ?CrowdingAtHome
WHERE { ?subjectOfcare fhir:Patient.forename ?forename.
 ?subjectOfcare   coc:Lay_off_from_job ?Lay_off_from_job.

OPTIONAL {?subjectOfcare
coc:Crowding_at_home ?CrowdingAtHome}
}
```

To understand the geospatial distribution of all patients and from where they belong (i.e. *CQ4*). we can run the `SPARQL` query as provided below:

```
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX shib: <http://www.semanticweb.org/shib>
PREFIX  coc: <http://purl.org/net/for-coc#>
PREFIX fhir: <http://hl7.org/fhir/>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
SELECT  DISTINCT    ?CountryOfBirth  ?prescribedQuantity
WHERE {    ?patientVisit shib:ageInYears ?Age.
            ?patientVisit shib:visitOfPatient ?patient.
            ?patient shib:roleOf ?person.
            ?person shib:forename ?PersonForename.
            ?person shib:surname ?PersonSurname.
OPTIONAL {?person shib:findingRelToBiologicalSex ?Sex.}
            ?person shib:countryOfBirth ?Country.
            ?Country rdfs:label ?CountryOfBirth.
            ?patient rdfs:label ?PatientUPI.
            ?patient shib:hasPrescription ?Prescription.
            ?Prescription rdfs:label ?PrescriptionID.
            ?Prescription shib:prescribedQuantity
            ?prescribedQuantity.
            ?Prescription shib:prescriptionDrug
            ?drugSubstance.
            ?drugSubtance rdf:type shib:drugOrMedicament.
            ?drugSubtance rdfs:label ?subtanceName.
            ?drugSubstance shib:legalDrugProduct
            ?drugProduct.
            ?drugProduct rdfs:label ?ProductName.
```

The supplementary document provided the results of CQ1 *Retrieve all patients who live in a low-income area* and CQ4 *Spatial Distribution of Prescribed Quantity for All Patients*, demonstrating how to use Google Chart functionality within the GraphDB SPARQL query panel. It shows that the prescribed quantity of specific drugs is higher for male patients than for female patients. Additional CQs and their SPARQL structures are provided in the *supplementary document* available online.

We conceptualize healthcare data interoperability not as a one-time achievement, but as an ongoing, dynamic process—particularly within the rapidly evolving landscape of Information and Communication Technologies (ICTs). In this paper, we present a substantial advancement in tackling the persistent challenges of interoperability by leveraging and adapting existing models and methodologies within the framework of the Linked Data paradigm. Our approach aims to develop a highly interconnected and semantically rich knowledge graph that enables seamless data integration and exchange across heterogeneous healthcare systems.

D. Case Study: SDH-Aware Personalized Care Planning

A key component of our solution is the use of GraphDB [16] and its advanced pattern-matching and semantic querying capabilities, which play a central role in constructing, maintaining, and exploiting the personalized knowledge graph. As illustrated in Fig. 5, this approach is demonstrated through the modeling of an individualized care

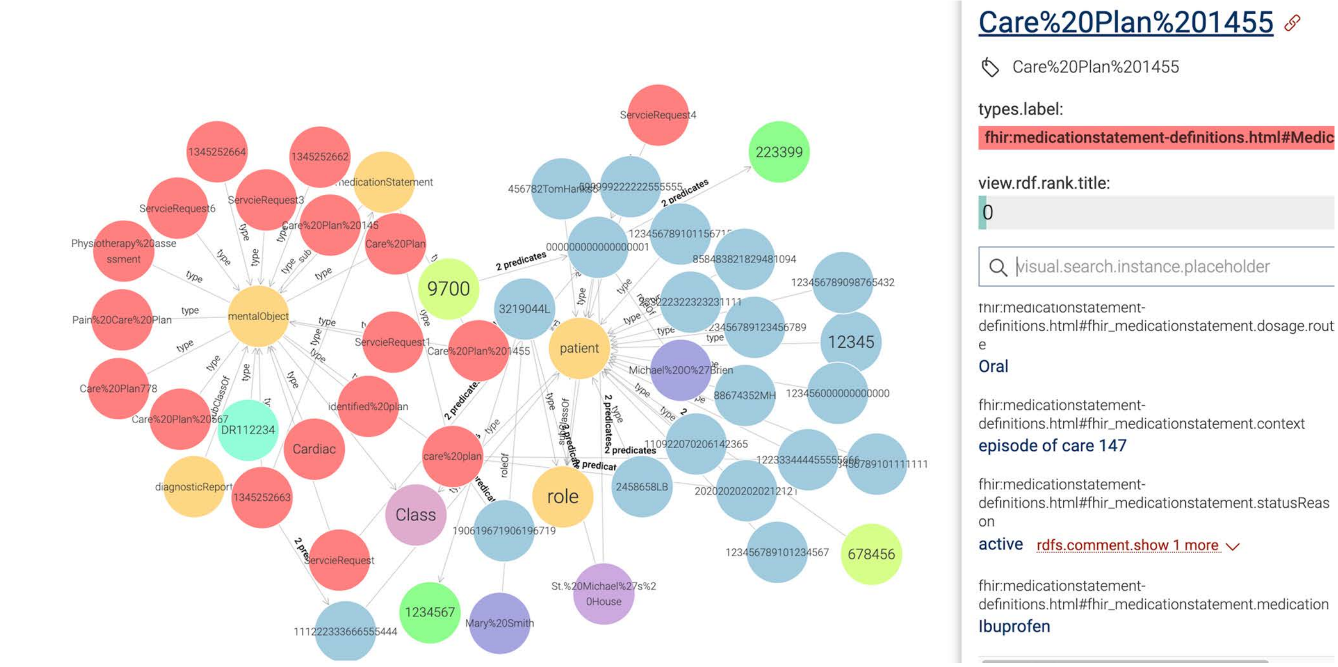


Fig. 5. CSSDH Knowledge Graph depicted specific care plan of Mary Smith.

plan for a patient identified as Mary Smith (ID: 3219044L). The care plan is explicitly linked to a specific episode of care (Episode 147), ensuring temporal and contextual coherence within the continuity-of-care framework. The personalized knowledge graph integrates both clinical and contextual information, including an active medication—Ibuprofen—together with its prescribed dosage and administration route (oral). By representing these elements as interconnected semantic entities, the knowledge graph enables a holistic and machine-interpretable view of the patient's situation. This structured representation supports automated reasoning, complex querying, and the dynamic enrichment of patient profiles, thereby facilitating personalized decision support and SDH-aware care planning. In a traditional EHR, these SDH factors would be recorded as unstructured notes or isolated fields. In CSSDH, they are formally linked to episodes of care and reasoning mechanisms, enabling automated queries and inference across clinical and social dimensions.

By utilizing knowledge graphs in this context, we significantly enhance the ability to manage and query complex healthcare data. Unlike traditional relational database systems that rely on costly and rigid join operations, knowledge graphs provide a more efficient and flexible means of handling intricate queries, supporting advanced analytics and personalized patient care.

The most recent version of the CSSDH OWL ontology is now accessible on BioPortal under the ContSOnto repository. In our ongoing research, we aim to assess user satisfaction by incorporating key user experience (UX) dimensions identified in prior studies on the Semantic User Interface (SemUI) [30]. This evaluation will help us refine the system's usability and effectiveness.

Looking ahead, we plan to validate our model using extensive real-world datasets. Specifically, we will conduct testing with the publicly available MIMIC-IV clinical dataset [31], which contains comprehensive electronic health records, as well as a regional dataset sourced from Hospital Recoletas Salud Zamora in Spain. This dual-dataset approach will allow for a thorough assessment of the model's generalizability across different healthcare environments.

In the final phase of our project, we will deploy the system on a cloud server in collaboration with our industry partner, *Davra.* This deployment will facilitate large-scale accessibility and real-time processing capabilities. Additionally, we will conduct rigorous performance testing of the CSSDH platform to ensure scalability, reliability, and efficiency in handling clinical data.

CRediT Authorship Contribution Statement

S. Das: developed the study concept, Methodology, and wrote the original draft.

D. Naskar: conducted the data collection.

S. Rodríguez González: provided feedback on the data and interpreted the results.

All authors contributed to drafting the manuscript, revised it, and approved the final version submitted for publication.

Data Statement

The primary datasets generated and analyzed during the current study are publicly available in the GitHub Repository repository at https://github.com/subhashishhh/CSDM. Additional data that support the findings are available from the corresponding author upon reasonable request. Additional CQs and their SPARQL structures are provided in the *supplementary document* available online.

Declaration of Conflicts of Interest

The authors declares that they have no financial benefit associated with this research work and have no conflict of interest to disclose.

Declaration of Generative AI and AI-Assisted Technologies

During the preparation of this work, the authors used ChatGPT-5.2 to check and improve the English grammar and sentence structure across all sections of the manuscript. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the final contents of the article.

Acknowledgment

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 101034371. The author is indebted to thank his colleagues from the National Standards Authority of Ireland (NSAI)-HISC, International Organization for Standardization (ISO)/TC215, and European Committee for Standardization (CEN)/TC251 for their kind and constructive support and cooperation. The project was approved by the President of the Research Ethics of University of Salamanca in Salamanca November, 2023 with document no InformeCEI_EC_en_715.

References

[1] Health Organization, "Integrating rehabilitation into health systems," 2024. Accessed: 2025-05-30. [Online]. Available: https://www.who.int/activities/integrating-rehabilitation-into-health-systems.

[2] World Health Organization, *Operational framework for monitoring social determinants of health equity*. World Health Organization, 2024.

[3] U.S. Department of Health and Human Services., "Healthy people 2030 building a healthier future for all," 2024. Accessed: 2025-05-30. [Online]. Available: https://odphp.health.gov/healthypeople

[4] S. Das, D. Naskar, S. R. González, "Cssdh: An ontology for social determinants of health to operational continuity of care data interoperability," in *International Conference on Intelligent Data Engineering and Automated Learning*, 2024, pp. 103–109, Springer. https://doi. org/10.1007/978-3-031-77731-8_10

[5] F. Pecoraro, D. Luzi, E. Pourabbas, F. L. Ricci, "A system of concepts to support the integration of health and social care and assistive domotics services: The health@ home project," *Informatics for Health and Social Care*, vol. 46, no. 3, pp. 333–344, 2021, doi: https://doi.org/10.1080/17538157.2021.1895167

[6] M. N. Cantor, L. Thorpe, "Integrating data on social determinants of health into electronic health records," *Health Affairs*, vol. 37, no. 4, pp. 585–590, 2018, doi: https://doi.org/10.1377/hlthaff.2017.1252

[7] N. F. Noy, N. H. Shah, P. L. Whetzel, B. Dai, M. Dorf, N. Griffith, C. Jonquet, D. L. Rubin, M.-A. Storey, C. G. Chute, *et al.*, "Bioportal: ontologies and integrated data resources at the click of a mouse," *Nucleic acids research*, vol. 37, no. suppl_2, pp. W170–W173, 2009, doi: https://doi.org/10.1093/nar/gkp440

[8] N. M. Kollapally, Y. Chen, J. Xu, J. Geller, "An ontology for the social determinants of health domain," in *2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2022, pp. 2403–2410, IEEE. https://doi.org/10.1109/BIBM55620.2022.9995544

[9] A. Jani, H. Liyanage, C. Okusi, J. Sherlock, U. Hoang, F. Ferreira, I. Yonova, S. de Lusignan, "Using an ontology to facilitate more accurate coding of social prescriptions addressing social determinants of health: feasibility study," *Journal of Medical Internet Research*, vol. 22, no. 12, p. e23721, 2020, doi: https://doi.org/10.2196/23721

[10] Y. Dang, F. Li, X. Hu, V. K. Keloth, M. Zhang, S. Fu, M. F. Amith, J. W. Fan, J. Du, E. Yu, *et al.*, "Systematic design and data-driven evaluation of social determinants of health ontology (sdoho)," *Journal of the American Medical Informatics Association*, vol. 30, no. 9, pp. 1465–1473, 2023, doi: https://doi.org/10.1093/jamia/ocad096

[11] D. Meyer, E. Lerner, A. Phillips, K. Zumwalt, "Universal screening of social determinants of health at a large us academic medical center, 2018," *American journal of public health*, vol. 110, no. S2, pp. S219–S221, 2020, doi: https://doi.org/10.2105/AJPH.2020.305747

[12] S. Das, P. Hussey, "Contsonto: a formal ontology for continuity of care," in *pHealth 2021*, IOS Press, 2021, p. 82–87, doi: https://doi.org/10.3233/shti210577

[13] International Organization for Standardization, "ISO/CD 13940 Health informatics — System of concepts to support continuity of care," 2024. Accessed: 2025-05-30. [Online]. Available: https://www.iso.org/standard/91284.html

[14] N. Guarino, *Formal ontology in information systems: Proceedings of the first international conference (FOIS'98), June 6-8, Trento, Italy*, vol. 46. IOS press, 1998. https://dl.acm.org/doi/10.5555/521669

[15] S. Das, P. Hussey, "Hl7-fhir-based contsys formal ontology for enabling continuity of care data interoperability," *Journal of Personalized Medicine*, vol. 13, no. 7, p. 1024, 2023, doi: https://doi.org/10.3390/jpm13071024

[16] GraphDB, "Graphdb workbench," 2024.Accessed: 2025-05-30. [Online]. Available: http://graphdb.ontotext.com.

[17] HL7, "Fhir release 4," 2024. [Online]. Available: https://hl7.org/FHIR/resourcelist.html

[18] N. Guarino, C. A. Welty, "An overview of ontoclean," *Handbook on ontologies*, pp. 201–220, 2009. https://doi.org/10.1007/978-3-540-92673-3_9

[19] C. Masolo, S. Borgo, A. Gangemi, N. Guarino, A. Oltramari, "Wonderweb deliverable d17: The wonderweb library of foundational ontologies," Laboratory for Applied Ontology (LOA), ISTC-CNR, 2002. [Online]. Available: https://cnr.it

[20] International Organization for Standardization, "ISO/IEC 21838-3:2023 Information technology—Top-level ontologies (TLO) Part 3: Descriptive ontology for linguistic and cognitive engineering (DOLCE)," 2024. Accessed: 2025-05-30. [Online]. Available: https://www.iso.org/standard/78927.html

[21] H. Stenzhorn, S. Schulz, M. Boeker, B. Smith, "Adapting clinical ontologies in real-world environments," *Journal of universal computer science*, vol. 14, no. 22, p. 3767, 2008, doi: https://doi.org/10.3217/jucs-014-22-3767

[22] C. Martínez-Costa, S. Kay, N. Oughtibridge, S. Schulz, "Contsys under ontological scrutiny.," in *MIE*, 2015, p. 999. https://doi.org/10.3233/978-1-61499-512-8-999

[23] N. Guarino, "The role of identity conditions in ontology design," in *Spatial Information Theory. Cognitive and Computational Foundations of Geographic Information Science: International Conference COSIT'99 Stade, Germany, August 25–29, 1999 Proceedings 4*, 1999, pp. 221–234, Springer. https://doi.org/10.1007/3-540-48384-5_15

[24] J. H. Gennari, M. A. Musen, R. W. Fergerson, W. E. Grosso, M. Crubézy, H. Eriksson, N. F. Noy, S. W. Tu, "The evolution of protégé: an environment for knowledge-based systems development," *International Journal of Human-computer studies*, vol. 58, no. 1, pp. 89–123, 2003, doi: https://doi.org/10.1016/S1071-5819(02)00127-1

[25] M. Grüninger, M. S. Fox, "The role of competency questions in enterprise engineering," in *Benchmarking—Theory and practice*, Springer, 1995, pp. 22–31, doi: https://doi.org/10.1007/978-0-387-34847-6_3

[26] T. Benson, G. Grieve, "Principles of health interoperability," *Cham: Springer International*, pp. 21–40, 2021, doi: https://doi.org/10.1007/978-3-030-56883-2

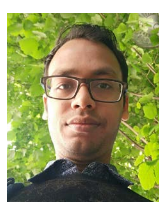
[27] F. Giunchiglia, S. Bocca, M. Fumagalli, M. Bagchi, A. Zamboni, "Popularity driven data integration," in *Iberoamerican Knowledge Graphs and Semantic Web Conference*, 2022, pp. 277–284, Springer. https://doi.org/10.1007/978-3-031-21422-6_20

[28] F. Baader, *The description logic handbook: Theory, implementation and applications*. Cambridge university press, 2003. https://doi.org/10.1017/CBO9780511711787

[29] B. Glimm, I. Horrocks, B. Motik, G. Stoilos, Z. Wang, "Hermit: an owl 2 reasoner," *Journal of automated reasoning*, vol. 53, pp. 245–269, 2014, doi: https://doi.org/10.1007/s10817-014-9305-1

[30] F. Giunchiglia, S. R. Ojha, S. Das, "Semui: a knowledge driven visualization of diversified data," in *2017 IEEE 11th International Conference on Semantic Computing (ICSC)*, 2017, pp. 234–241, IEEE. https://doi.org/10.1109/ICSC.2017.53

[31] A. Johnson, L. Bulgarelli, T. Pollard, S. Horng, L. A. Celi, R. Mark, "Mimic-iv," *PhysioNet. Available online at: https://physionet. org/content/mimiciv/1.0/(accessed August 23, 2021)*, pp. 49–55, 2020, doi: https://doi.org/10.13026/kpb9-mt58



Subhashis Das

Subhashis is a USAL4EXCELLENCE Marie Skłodowska-Curie postdoctoral fellow in the BISITE Research Group at the Universidad de Salamanca, Spain. He was an ELITE-S MSCA fellow at the ADAPT Centre, Dublin City University (2020–2023), and a member of the National Standards Authority of Ireland. He obtained his Ph.D. from the DISI, Università degli Studi di Trento, Italy, and completed an internship at the School of Informatics, University of Edinburgh, UK.

Debashis Naskar

Debashis is a USAL4EXCELLENCE Marie Skłodowska-Curie postdoctoral fellow in the BISITE Research Group at the Universidad de Salamanca, Spain. He received his Ph.D. in Computer Science from the Universitat Politècnica de València, Spain. His research interests include social network analytics, sentiment analysis and emotion detection.

Sara Rodríguez González

Sara Rodríguez González is a Full Professor in the Department of Computer Science and Automation at the Universidad de Salamanca, Spain. She earned her Ph.D. in Computer Science from the same university in 2010.