# Database Federation for Spatial Digital Twins with Semantically Enhanced Data Processing: A use case of Mission Mjøsa

Sajith Ranatunga[1], Rune Strand Ødegård[1], Erling Onstein[1], Knut Jetlund[1,2]

[1] NTNU Gjøvik, Teknologiveien 22, 2815 Gjøvik, Norway - sajith.ranatunga@ntnu.no; rune.oedegaard@ntnu.no; erling.onstein@ntnu.no;
[2] Kartverket, Kartverksveien 21, 3511 Hønefoss, Norway - knut.jetlund@ntnu.no

**Keywords:** Semantic Web Technologies, OBDA, SPARQL, Database Federation, Spatial Digital Twin

**Abstract**

Developing spatial digital twins (SDTs) for inland water bodies requires addressing challenges such as temporal variability, dynamic water surface conditions, and heterogeneous water datasets. This study presents a semantic integration framework combining Ontology-Based Data Access (OBDA) with PostgreSQL database federation to support real-time decision-making in SDT applications. The framework integrates hydrological data, sensor networks, bathymetric data, and other static datasets under a unified ontology, enabling federated SPARQL queries without centralizing data. A 3D visualization interface allows interactive analysis of results from the SPARQL queries. Demonstrated through the Mission Mjøsa use case, this approach supports inland water mapping and decision support by aligning spatial data across domains. The framework offers a scalable solution for water data processing, semantic fusion, and scenario-based simulation, contributing to the development of intelligent, semantically enriched SDTs for sustainable environment management.

## 1. Introduction

Water bodies such as lakes, rivers, and reservoirs are critical ecological and socio-economic assets that provide essential services, including freshwater supply, biodiversity support, recreation, flood regulation, and climate stabilization. Monitoring and managing these dynamic aquatic environments require robust data integration strategies capable of bringing together heterogeneous datasets from diverse domains, such as hydrology, meteorology, land use, and environmental monitoring. In this context, the emergence of spatial digital twins (SDTs) offers a transformative approach to visualize, simulate, and manage water systems by replicating their real-world behaviours in a digital environment, and they are also designed to capture the spatial context of geospatial objects, including their location and dimensional relationships (Ali et al., 2023).

Unlike traditional geographic information systems (GIS), SDTs offer continuous, real-time feedback loops that mirror changes in the physical environment through integration with sensor networks, hydrological data, and geospatial databases. This is particularly important for water bodies like **Lake Mjøsa** (Figure 1), Norway's largest inland lake, where decision-makers must respond to rapid changes in water levels, pollution events, and climate-induced variability. However, integrating these various datasets remains a significant challenge due to differences in data formats, temporal resolution, spatial reference systems, and domain-specific semantics.

Historically, water-related data management has relied on centralized GIS platforms and relational databases. While effective for visualizing structured geospatial data, these systems struggle to efficiently integrate and query dynamic, multi-source datasets. Challenges include data fragmentation across institutional silos, inconsistent schemas, and limited capabilities for semantic querying (Dao et al., 2024; Loukili et al., 2022). These issues hinder real-time interoperability and constrain the potential of SDTs in environmental monitoring, flood risk management, and policy decision-making.

To overcome these challenges, recent developments in semantic web technologies including Ontology-Based Data Access (OBDA) offer promising solutions. OBDA introduces a virtual semantic layer over relational databases by mapping them to domain ontologies, allowing users to write intuitive, high-level SPARQL queries without needing to transform or replicate the underlying data (Allemang et al., 2020; Kharlamov et al., 2017; Xiao et al., 2018). Tools like Ontop automate the translation of SPARQL to SQL, enabling semantically rich queries on federated datasets (Calvanese et al., 2016).

Figure 1. 3D view of Lake Mjøsa and surrounding area

A key feature of OBDA is its ability to connect relational databases with a semantic layer, enabling seamless data integration and querying through shared vocabularies and conceptual models. This approach is particularly powerful in interdisciplinary domains like hydrology and water resource management, where domain experts may not be familiar with database schemas but can query for meaningful concepts like “flood zone,” “pollution event,” or “riparian vegetation.” Additionally, Linked Data principles allow disparate data elements to be semantically connected, enhancing interoperability and data reusability (Allemang et al., 2020).

To enable scalable and real-time access to distributed datasets, OBDA can be combined with database federation techniques such as PostgreSQL Foreign Data Wrappers (FDWs). This allows users to perform live SPARQL queries over multiple databases located across institutions, while maintaining data autonomy and avoiding data duplication (Wen et al., 2022).

Database federation plays a pivotal role in water domain applications, where data is often collected and maintained by different authorities such as water utilities, environmental agencies, and meteorological services. Using FDWs, the federated infrastructure allows for on-the-fly access to these sources, enabling SDTs to perform integrated analyses across datasets such as water level logs, rainfall records, land cover classifications, and pollution measurements. Semantic integration through OBDA enhances this capability by aligning diverse schemas under a shared conceptual model, enabling the exploration of complex environmental relationships (e.g., the impact of rainfall on nutrient runoff or the correlation between land use and water quality).

For example, a SDT of a river/lake system may integrate topographic maps, rainfall data, in-situ sensor readings, and citizen science observations using a shared semantic model. This enables more advanced analyses, such as identifying correlations between rainfall intensity and turbidity, detecting pollution hotspots, or visualizing the spatiotemporal spread of sediment. The integration of OBDA and FDWs thus creates a flexible and adaptive framework suitable for SDTs of aquatic environments.

By combining semantic data processing with federated access and OBDA, SDTs can achieve a high degree of flexibility, adaptability, and intelligence which are critical features for managing complex and dynamic systems like lakes, rivers, and wetlands under changing environmental conditions.

The primary objective of this study is to demonstrate the effectiveness of a federated OBDA framework for integrating and querying heterogeneous water-related datasets in support of a SDT of Lake Mjøsa. By implementing a federated data architecture enriched with ontologies and linked data principles, the proposed framework allows for semantically enhanced queries that support real-time environmental monitoring and analysis. It is essential to emphasize their application in solving real-world challenges associated with the management and monitoring of freshwater systems. In the context of Mission Mjøsa, these technologies are used not merely for generic geospatial data handling but specifically to support hydrological and hydraulic applications, including semantic integration of multi-source water data, lake dynamics analysis, and flood risk management.

To demonstrate the practical application of the proposed framework, Mission Mjøsa (*Oppdrag Mjøsa - NTNU*, 2025) serves as an illustrative use case. As Norway's largest lake, Mjøsa is vital to many communities but faces increasing challenges from development, industry, agriculture, and climate change. In response, Innlandet County Municipality, regional municipalities, and NTNU have launched Mission Mjøsa, an SDT initiative aimed at improving environmental monitoring and management. By integrating data from multiple sources, this initiative enhances sustainability and value creation in the region through coordinated strategies and stakeholder collaboration.

Water bodies such as Lake Mjøsa present unique challenges for SDTs due to their dynamic and often non-static nature. The lake's depth, shoreline contours, inflow and outflow rates, sediment transport, and seasonal fluctuations in water quality all introduce a high degree of temporal and spatial complexity. These variations require continuous monitoring and high-resolution, multi-source data integration to create accurate digital representations and predictive models.

This research addresses the central question: *How can the combination of database federation using Foreign Data Wrappers (FDW) and Ontology-Based Data Access (OBDA) enhance semantically enriched data processing for SDT applications?* By integrating these two approaches, the paper aims to provide a framework that not only unifies disparate data sources but also enriches them with domain semantics, thereby improving near real-time decision-making capabilities in SDTs.

This paper presents a practical implementation of this framework through a custom-built web-based platform, integrating spatial and environmental datasets relevant to Lake Mjøsa. The system demonstrates how semantic technologies, in combination with database federation, can facilitate timely, informed decision-making in domains such as flood forecasting, pollution tracking, and aquatic ecosystem management. The platform further includes a 3D visualization environment to support scenario exploration and stakeholder engagement, essential for mission-driven environmental initiatives.

The remainder of this paper is organized as follows. Section 2 describes the methodology and structure of the proposed framework, including ontology development and federated query processing. Section 3 presents the implementation details, and a use case focused on Lake Mjøsa. Section 4 discusses the system's advantages, limitations, and implications for water resource management. Finally, Section 5 concludes the paper with future research directions.

## 2. Methodology

### 2.1 Overview of the framework

This section outlines the methodology for developing and implementing an OBDA framework for spatial data analysis. The framework integrates heterogeneous datasets, specifically official data from the Norway geodata portal (geonorge.no), to enable advanced querying and analysis using semantic web technologies. The framework consists of three main components: ontology development and mapping, SPARQL query formulation and execution, and presenting the results in a custom-built web-based 3D environment, which lays a foundation for the future semantically rich SDTs in spatial data management.

The data from GeoNorge.no was downloaded and subsequently stored in various PostgreSQL relational databases to mimic the distributed data sources and enhanced with the PostGIS extension to support geospatial functionality. This setup serves as the central data store and external data sources, enabling seamless integration with the OBDA system for efficient querying and analysis of geospatial information.

### 2.2 Ontology Development and Mapping

It is best practice to reuse and extend existing standard ontologies (Noy & McGuinness, 2001) such as GeoSPARQL, SSN (Semantic Sensor Network Ontology), and SOSA (Sensor, Observation, Sample, and Actuator Ontology). This ensures compatibility and enhances interoperability with other datasets and systems. For this specific ontology development, well-known GeoSPARQL ontology by the Open Geospatial Consortium[1] was used and extended to be compatible with demonstrating use case.

A simple ontology based on GeoSPARQL for the experimental use case was developed. It can be modelled with various aspects water related data, including water level, water temperature, water flow and their relationships to geographic and environmental features. The ontology serves as a unified conceptual model that abstracts the underlying data sources and facilitates interoperability between them. As shown in Figure 2, several classes and data properties were added to the base GeoSPARQL ontology within the Protégé[2] ontology editing application.

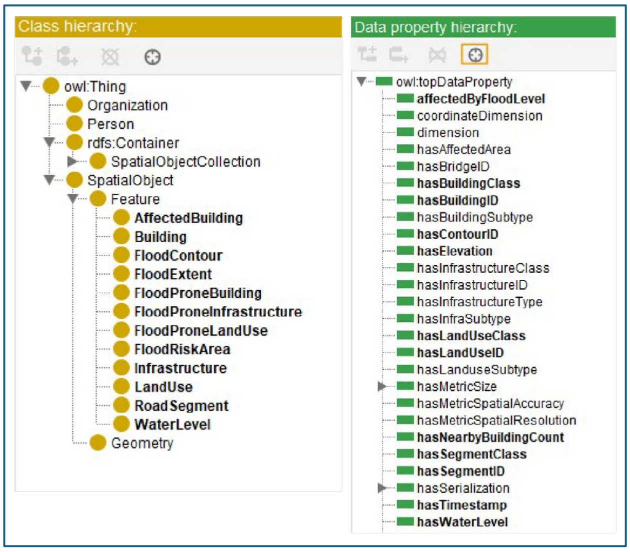


Figure 2. Class and Data Property hierarchy for a custom ontology for use cases

To effectively integrate and utilize these datasets within a unified analytical framework, it is essential to carefully map them to a shared ontology that captures the semantic relationships between these different types of information. For this task, the Ontop Protégé Plugin was used. Figure 3 shows an example mapping for water level data within the Protégé.

Ontop[3], an open-source OBDA platform, plays a central role in this framework by providing the tools necessary for creating semantic mappings between the relational data sources (the Geonorge and other static datasets) and the ontology that represents the geospatial domain. Ontop enables the seamless integration of these datasets by using mappings to translate SPARQL queries, native to the semantic web, into SQL queries that can be executed directly on the relational database where the data is stored. This process eliminates the need for physical data transformation or restructuring, allowing the datasets to remain in their original formats while still being accessed through a unified, ontology-driven interface. The basis for this framework was based on the Ontop base system explained by (Calvanese et al., 2016).

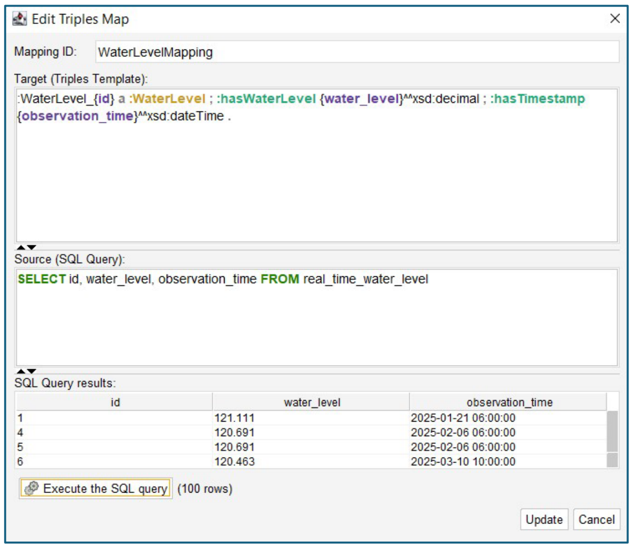


Figure 3. Example mapping for near real-time water level

Using Ontop's capabilities, the OBDA framework ensures that data can be queried and analysed cohesively. Ontop facilitates the alignment of different aspects of water data within a common semantic framework, enabling comprehensive analyses of the spatial relationships and interactions among water feature characteristics. This approach enhances interoperability, supports advanced geospatial queries, and maximizes the value of the integrated datasets by allowing users to derive meaningful insights from the combined information.

### 2.3 Database federation using PostgreSQL

To enable seamless access to distributed datasets, database federation was implemented using PostgreSQL's FDWs. FDWs allow one PostgreSQL instance (central database) to query and interact with other external PostgreSQL or PostGIS databases without physically replicating the data. In this setup, a central database serves as the integration layer, connecting to multiple external sources based on a local network containing hydrological, meteorological, and spatial datasets. This approach ensures that each data source remains autonomous while supporting unified query execution across systems. The federated structure simplifies real-time data access, enhances modularity, and enables efficient integration of heterogeneous water-related datasets for the SDTs.

### 2.4 SPARQL Query Formulation and Execution

The OBDA framework utilizes SPARQL, enhanced with GeoSPARQL functions, to execute advanced geospatial queries across integrated datasets. This enables users to perform spatial analyses directly on federated data sources, such as those from GeoNorge.no. Beyond basic data retrieval, the use of SPARQL allows for querying complex spatial relationships such as containment, intersection, and proximity between geospatial features from multiple datasets. This capability is essential for extracting meaningful insights in environmental monitoring and water resource management, where spatial patterns and interactions play a critical role.

[1] https://www.ogc.org/standard/geosparql/
[2] https://protege.stanford.edu/
[3] https://ontop-vkg.org/


https://doi.org/10.5194/isprs-archives-XLVIII-2-W10-2025-231-2025  

### 2.5 The Proposed Framework

The proposed OBDA framework integrates with a custom-built web-based 3D visualization environment using HTML, CSS, and JavaScript to enhance the visualization and user engagement of SPARQL query results. This interactive platform dynamically displays geospatial data, allowing users to explore data visually. By using Mapbox web maps, the 3D visualization environment can be used to render spatial relationships and features, which provide a more intuitive and immersive understanding of the data. This approach improves data accessibility, improves the decision-making process, and lays the groundwork for developing semantically rich SDTs, where real-world environments are modelled in digital space, facilitating advanced simulations, planning, and analysis.

The proposed framework integrates FDW-driven database federation with OBDA to enhance data processing for SDTs. It is structured into four key layers. Figure 4 shows hierarchical structure of the framework.

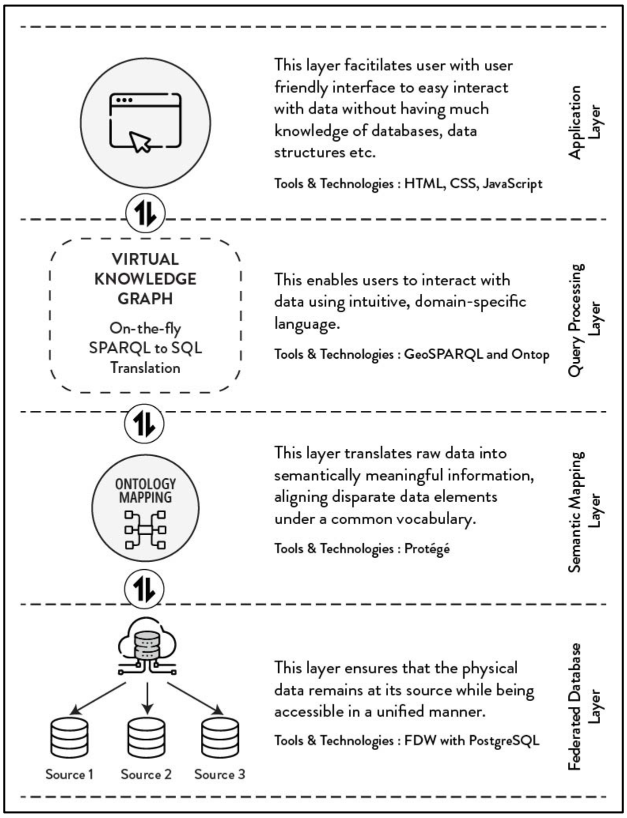


Figure 4. Hierarchical structure of the proposed framework

The federated database layer utilizes FDW to connect and query distributed databases. The semantic mapping layer applies domain-specific ontologies to the federated data. Then the query processing layer uses OBDA to convert high-level SPARQL queries into optimized SQL queries that the federated database layer can execute. Finally, the application layer interfaces with the SDT platform or module of a large SDT application, delivering enriched data for real-time monitoring, analysis, and decision-making.

As a proof of concept, this platform was tested in a local environment using PostgreSQL and the PostGIS extension to handle spatial data. To simulate a federated database system, three databases were created: one serving as the central database and the other two mimicking external data sources. The central database was designed to aggregate and integrate data from these external sources dynamically. Additionally, a dedicated table was implemented in the central database to store real-time water level readings of the lake on demand, updating the application was executed. The external databases were seamlessly connected to the central database using FDWs, enabling cross-database queries without requiring data duplication. This setup effectively demonstrated the feasibility of integrating distributed geospatial data in real-time while ensuring efficient data access and management.

The integrated platform connects with the central database using the Ontop plugin, enabling semantic querying of spatial and non-spatial data stored in the federated PostgreSQL database. Ontop acts as a virtual layer, mapping relational data to an ontology, allowing users to retrieve information using SPARQL without modifying the underlying database. Through predefined mappings, the central PostgreSQL database, aggregating critical environmental datasets, and external data sources, such as real-time hydrological observations (temperature, water level, water flow) dynamically retrieved via Application Programming Interfaces (APIs), are federated using FDWs and exposed as a unified knowledge graph. This integration ensures seamless access to heterogeneous geospatial datasets while maintaining interoperability and real-time data retrieval capabilities.

## 3. Demonstration: Use Case of Mission Mjøsa

### 3.1 Overview of the use cases

The custom-built web-based application is developed to demonstrate the practical implementation of the proposed system. This system integrates data from multiple sources, enabling a comprehensive and dynamic environmental monitoring approach. The implementation of Mission Mjøsa highlights that this integrated approach not only enhances the efficiency and accuracy of data processing but also strengthens the SDT's ability to provide real-time, semantically enriched information. This, in turn, facilitates timely decision-making, particularly in critical scenarios such as flood management, ultimately supporting more effective environmental management and policy development. Figure 5 shows the user interface of custom-built web application with sample query to retrieve data. It also shows the result of applied SPARQL query (Listing 1).

### 3.2 Use Case: Identifying Critical Infrastructure within Flood Risk Zones

Objective: The objective of this use case is to identify critical infrastructure buildings, specifically hospitals and schools, that are located within designated flood risk areas in the Lake Mjøsa region. This analysis supports emergency planning, infrastructure protection, and risk mitigation efforts by enabling authorities to quickly determine which high-priority facilities are vulnerable to flooding.

Use: This use case utilizes a federated OBDA system that integrates hydrological and infrastructure datasets through a semantic ontology-based model. Using GeoSPARQL-enabled SPARQL queries, it performs a spatial containment analysis between building footprints and mapped flood risk zones. The query:

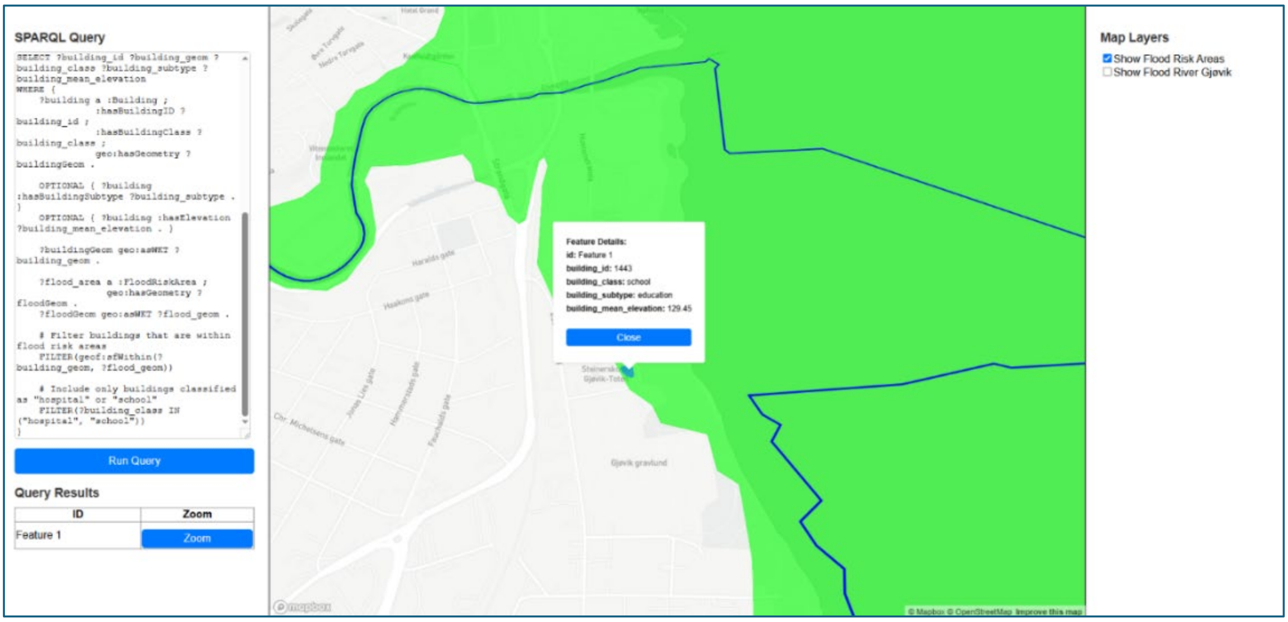

Figure 5. SPARQL query user interface, executed query, and the result

- Retrieves buildings classified as hospitals or schools from the building dataset.
- Filters those buildings that are spatially within the geometry of a flood risk area.
- Returns the building's ID, geometry, type, subtype (if available), and mean elevation (if recorded), using semantic mappings defined in the OBDA framework.

The building data is retrieved from a spatial database (e.g., PostGIS) accessed through PostgreSQL FDW, while flood risk areas are sourced from hydrological models or hazard maps published through national platforms such as GeoNorge. All data is semantically aligned through the predefined ontology developed for this SDT application. Listing 1 presents the example SPARQL query.

Listing 1. Example SPARQL Query for use case 1 (Q1)

```
SELECT ?building_id ?building_geom ?building_class
?building_subtype ?building_mean_elevation
WHERE {
    ?building a :Building ;
              :hasBuildingID ?building_id ;
              :hasBuildingClass ?building_class ;
              geo:hasGeometry ?buildingGeom .

    OPTIONAL { ?building :hasBuildingSubtype
?building_subtype . }
    OPTIONAL { ?building :hasElevation
?building_mean_elevation . }

    ?buildingGeom geo:asWKT ?building_geom .

    ?flood_area a :FloodRiskArea ;
                geo:hasGeometry ?floodGeom .
    ?floodGeom geo:asWKT ?flood_geom .

    # Filter buildings that are within flood risk areas
    FILTER(geof:sfWithin(?building_geom, ?flood_geom))

    # Include only buildings classified as "hospital"
or "school"
    FILTER(?building_class IN ("hospital", "school"))
}
```

Result:
The query returns a list of all hospitals and schools located within flood-prone zones of the Lake Mjøsa catchment, including their spatial geometries and elevation attributes. This result allows stakeholders such as municipal planners, emergency response teams, and environmental agencies—to:

- Visualize at-risk buildings within the SDT interface.
- Assess the elevation profile of vulnerable facilities to determine likely flood impact.
- Prioritize infrastructure for evacuation planning, floodproofing, or future relocation.

The semantically enriched output can also be used to generate alerts, simulate response scenarios in a 3D environment, or be integrated into broader risk assessment frameworks. By linking hydrological and built environment data through OBDA and federated access, this use case demonstrates how semantic technologies can support targeted, high-value decision-making in real-time water risk contexts.

## 4. Discussion

SDTs for water bodies, such as lakes, rivers, and coastal zones, introduce a distinct set of challenges that differ from those of built environments. One critical aspect is the temporal variability of water systems, which are influenced by seasonal changes, weather events, snowmelt, and long-term climate trends. Fluctuations in water levels, flow rates, and thermal stratification can significantly affect aquatic ecosystems and surrounding communities.

In addition to temporal dynamics, the diversity of data types needed to model water bodies are particularly complex. Effective monitoring requires integrating data from multiple domains, including hydrological models, sensor networks (e.g., for turbidity, temperature, pH), remote sensing imagery, and bathymetric surveys. These data sources vary in structure, frequency, and spatial resolution, creating substantial barriers to interoperability and analysis. Water systems are also highly environmentally dependent. Their conditions are shaped by a combination of natural factors (e.g., precipitation, runoff) and human activities (e.g., agriculture, wastewater discharge, land use changes). This interdependency demands an integrated approach that links datasets from various domains, including climate, urban planning, ecology, and public health.

Finally, water bodies often require real-time decision-making to manage events such as floods, pollution spills, and harmful algal blooms. This necessitates a system that not only integrates data from multiple sources but also supports fast, context-aware querying and interpretation.

The proposed framework addresses these challenges by combining database federation and OBDA. The semantic layer bridges cross-domain datasets by mapping them to a unified ontology, allowing high-level, concept-driven queries across hydraulic, meteorological, and environmental datasets. This enables decision-makers to perform holistic analyses, such as identifying correlations between rainfall intensity and pollution dispersion or linking satellite-derived land use changes to in-lake nutrient levels. By semantically enriching distributed data and enabling real-time querying, the framework equips SDTs of water bodies with the adaptability and intelligence needed to support sustainable and resilient water management.

As presented in Section 3, semantic technologies show strong potential for advancing the integration and analysis of water-related geospatial data. The proposed framework and its associated use case demonstrated the effective application of semantic web technologies in managing and analysing environmental data concerning freshwater systems. By constructing a shared ontology to represent diverse hydrological, environmental, and sensor-based datasets, the framework facilitates semantic interoperability and enables meaningful integration of heterogeneous information sources. The OBDA-based approach, powered by Ontop, dynamically maps relational databases to a semantic layer without requiring physical data transformation. This allows for sophisticated SPARQL querying, supporting complex environmental analyses such as pollutant tracing, water level fluctuation monitoring, or linking rainfall events to downstream turbidity spikes.

A notable strength of the framework is its capacity to unify datasets from varied origins. In this study, very few datasets are incorporated for the demonstrative implementation, but more relevant data to freshwater ecosystems such as water level readings, weather data, and water quality parameters can be integrated upon minor modifications to the ontology. These datasets can originate from sensor networks, meteorological stations, hydrological models, and public repositories. The ability to integrate and harmonize such diverse data streams is vital in water-related applications, where system behaviours are influenced by both natural events and anthropogenic pressures. Moreover, these datasets often update in near real-time, and their semantic alignment allows stakeholders to interpret, compare, and analyze the data within a unified conceptual framework.

The semantic integration of water-related data encourages stakeholders such as hydrologists, environmental authorities, municipalities, and emergency responders to access enriched datasets that reveal deeper insights into spatial and temporal water dynamics. For instance, the ability to perform federated SPARQL queries across multiple databases enables a user to ask, "What is the water quality in regions experiencing elevated rainfall this week?" or "Which catchment areas contribute to current nutrient spikes in the lake?" questions that would otherwise require extensive pre-processing and manual correlation. The integration of OBDA with PostgreSQL FDWs ensures that these queries can be executed on live datasets across distributed systems, improving both accessibility and responsiveness.

By enabling advanced spatial queries and harmonizing data from diverse domains, this framework can be developed to support a wide range of underwater decision-making applications. These include bathymetric change detection, sediment transport monitoring, submerged infrastructure assessment, and aquatic habitat mapping. The analytical results are visualized within a 3D geospatial interface, allowing users to interact in real time with underwater terrain models, lakebed morphology, and subsurface hydrological features, such as inflows, outflows, and stratification zones. Crucially, the framework avoids the need for centralized data storage by leveraging federated access, enabling efficient, on-demand querying of large and distributed datasets an essential capability for scalable, real-time monitoring of underwater environments.

However, as with any semantic integration system, challenges remain. Data quality is a consistent concern especially when integrating datasets from heterogeneous sources such as community sensors, legacy hydrological systems, or open environmental APIs. These sources may vary in granularity, accuracy, and update frequency, which can introduce inconsistencies or bias. As such, the development of robust validation and quality assurance pipelines is necessary to maintain the reliability of analytical outputs. Furthermore, large-scale geospatial queries especially those involving complex 3D geometries and temporal joins can be computationally intensive. Optimization techniques such as query rewriting, caching, and spatial indexing are recommended to maintain system performance and user responsiveness.

Another challenge lies in current limitations of the OBDA platforms. For example, Ontop does not yet fully support GeoSPARQL 1.1, which limits the scope of geospatial functions available directly through SPARQL. Consequently, certain spatial operations (e.g., proximity buffering, 3D intersection analysis) must be handled externally, reducing the semantic coherence of the analytical workflow. Furthermore, the framework requires specialized knowledge in semantic technologies, geospatial data processing, and environmental domain modelling skills that are not always readily available in water management institutions.

Despite these limitations, the OBDA-based framework with its 3D visualization interface offers a powerful foundation for developing SDTs for water bodies. It enables real-time semantic querying and integration of diverse hydrological, environmental, and geospatial datasets, providing a holistic understanding of aquatic systems. This is particularly relevant for Mission Mjøsa, where multiple stakeholders are concerned with dynamic lake monitoring, flood mitigation, ecological protection, and climate adaptation.

The 3D interface plays a pivotal role in enabling intuitive and exploratory interaction with complex hydrological systems. Users can visualize temporal water level changes, overlay predicted flood zones, and simulate various environmental management scenarios. With continuous data updates and semantic enrichment, this framework supports proactive decision-making and scenario modelling cornerstones of a fully functional SDT.

As SDTs continue to evolve, they will increasingly require real-time, semantically consistent, and federated access to water-related data. The presented approach addresses these requirements and provides a flexible, extensible foundation for future advancements. Future work may explore predictive analytics such as modelling nutrient runoff under various land use scenarios or simulating the ecological impact of prolonged droughts enabled by integrating machine learning models into the semantic framework.

In summary, this OBDA-based integration framework with semantic support and 3D visualization capabilities represents a robust and scalable solution for managing and analysing complex datasets related to water bodies. It advances the creation of intelligent, semantically rich, and future-ready SDTs for sustainable water resource management.

## 5. Conclusion

This research demonstrates that the integration of FDW-based database federation with OBDA provides a powerful framework for enhancing data processing in SDT applications. By addressing both structural and semantic challenges, the proposed approach improves real-time data integration, enhances query capabilities, and contributes significantly to the operational intelligence of SDTs. The Mission Mjøsa use case underscores the practical benefits of this methodology, offering a promising direction for future research and application in diverse domains.

The potential of semantic web technologies for advanced spatial analysis of geospatial data particularly for underwater environments remains significantly under-explored. Water bodies pose unique integration and analysis challenges due to the complexity of submerged terrain, variable hydrodynamics, and the need for high-resolution, multi-dimensional data. Several key improvements are suggested for future work to enhance the usability and applicability of the proposed framework for SDTs. First, expanding the ontology to incorporate underwater-specific classes such as bathymetric features, sediment layers, aquatic vegetation, and submerged infrastructure would provide a more detailed semantic understanding of lakebed dynamics and subsurface conditions.

Furthermore, incorporating database federation techniques enables real-time querying across distributed underwater monitoring systems (e.g., bathymetric scans, in-situ sensors, sonar data) without the need for centralized storage or redundant data movement. Enhancing the framework with user-friendly interfaces and interactive 3D visualizations such as dynamic lakebed models or sediment accumulation simulations will improve accessibility for a broader range of stakeholders, including environmental researchers, marine engineers, and emergency response planners.

Managing and querying 3D volumetric data such as depth profiles, terrain elevation, and water column structure poses significant challenges for accurate underwater SDTs. These data types often involve large volumes, complex geometries, and varying temporal resolutions, making them difficult to integrate and process within standard geospatial frameworks. However, by using OBDA, the proposed framework offers great potential to integrate these high-complexity data structures in a more accessible and semantically consistent way. Through ontology-driven mappings and federated access, it becomes possible to abstract and query these datasets using high-level concepts without directly handling the underlying technical intricacies.

Future extensions may integrate data types such as LiDAR bathymetry, sonar-derived point clouds, 3D terrain surfaces, and volumetric hydrological models, which would significantly improve the framework's capability to monitor and query underwater systems. Ultimately, applying this approach to create semantically enriched SDTs for aquatic environments will unlock new possibilities for real-time lakebed monitoring, habitat analysis, flood preparedness, and environmental protection supporting innovation in sustainable water resource management and SDT-driven decision support.

## References

Ali, M. E., Cheema, M. A., Hashem, T., Ulhaq, A., & Babar, M. A. (2023). *Enabling Spatial Digital Twins: Technologies, Challenges, and Future Research Directions* (No. arXiv:2306.06600; Version 1). arXiv. http://arxiv.org/abs/2306.06600

Allemang, D., Hendler, J., & Gandon, F. (2020). *Semantic Web for the Working Ontologist: Effective Modeling for Linked Data, RDFS, and OWL* (3rd ed.). ACM. https://doi.org/10.1145/3382097

Calvanese, D., Cogrel, B., Komla-Ebri, S., Kontchakov, R., Lanti, D., Rezk, M., Rodriguez-Muro, M., & Xiao, G. (2016). Ontop: Answering SPARQL queries over relational databases. *Semantic Web*, *8*(3), 471–487. https://doi.org/10.3233/SW-160217

Dao, J., Ng, S. T., & Kwok, C. Y. (2024). Interlinking BIM and GIS data for a semantic pedestrian network and applications in high-density cities. *Developments in the Built Environment*, *17*, 100367. https://doi.org/10.1016/j.dibe.2024.100367

Kharlamov, E., Hovland, D., Skjæveland, M. G., Bilidas, D., Jiménez-Ruiz, E., Xiao, G., Soylu, A., Lanti, D., Rezk, M., Zheleznyakov, D., Giese, M., Lie, H., Ioannidis, Y., Kotidis, Y., Koubarakis, M., & Waaler, A. (2017). Ontology Based Data Access in Statoil. *Journal of Web Semantics*, *44*, 3–36. https://doi.org/10.1016/j.websem.2017.05.005

Loukili, Y., Lakhrissi, Y., & Ali, S. E. B. (2022). Geospatial Big Data Platforms: A Comprehensive Review. *KN - Journal of Cartography and Geographic Information*, *72*(4), 293–308. https://doi.org/10.1007/s42489-022-00121-7

Noy, N. F., & McGuinness, D. L. (2001). *Ontology Development 101: A Guide to Creating Your First Ontology*. https://protege.stanford.edu/publications/ontology_development/ontology101.pdf

*Oppdrag Mjøsa—NTNU*. (2025). https://www.ntnu.no/oppdrag-mjosa

Wen, Y., Zhang, H., Liu, X., Zhou, Y., Gao, W., & Zhu, X. (2022). Federal database system based on foreign data wrapper. *2022 International Conference on Frontiers of Communications, Information System and Data Science (CISDS)*, 112–116. https://doi.org/10.1109/CISDS57597.2022.00025

Xiao, G., Calvanese, D., Kontchakov, R., Lembo, D., Poggi, A., Rosati, R., & Zakharyaschev, M. (2018). Ontology-Based Data Access: A Survey. *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence*, 5511–5519. https://doi.org/10.24963/ijcai.2018/777