Towards FAIR Declarative Mappings: Evaluating MetaSEMAP's Usability in Lifecycle Documentation and Metadata Annotation

Sarah Alzahrani[1] and Declan O'Sullivan[2]

[1] Computer Science and Statistics, Trinity College Dublin, Dublin, Ireland
salzahra@tcd.ie
[2] ADAPT Centre, Computer Science and Statistics, Trinity College Dublin, Dublin, Ireland
declan.osullivan@tcd.ie

Abstract. Achieving FAIR (Findable, Accessible, Interoperable, Reusable) principles for declarative mapping projects requires robust documentation and metadata annotation throughout their lifecycle. This paper presents a usability evaluation of MetaSEMAP, a tool designed to facilitate the annotation of mappings and enhance the reusability of declarative mappings: uplift mappings, ontology alignment, and interlinking. The evaluation explores the tool's ability to annotate mapping projects with metadata drawn from real-world scenarios like the Virtual Record Treasury of Ireland. Participants provided feedback on the tool's usability and preferred metadata representation, including RDF-star and Named Graph. The results highlight key strengths and areas for improvement for MetaSEMAP, contributing valuable insights to the development of more effective tools for metadata annotation. This work lays the groundwork for future enhancements in mapping lifecycle documentation, aiming for interoperability and sustainability aligned with FAIR principles.

Keywords: FAIR principles · Metadata annotation · Mapping lifecycle · RDF-star · Named Graph

1 Introduction

Due to semantic heterogeneity and diverse data, there are constraints on interoperability among various web-based information systems [1]. This challenge becomes apparent when dealing with disparate datasets, ontologies, and the publication of data on the web [2]. To address this limitation, declarative mappings have been proposed as an effective approach for articulating the semantic relationships between two data models [3, 4]. Such declarative mappings are deployed in a range of diverse areas, from real time event distribution systems [5] to domain applications such as Climate Change applications [6].

Despite the potential of declarative mappings, researchers face numerous challenges in locating, accessing, comprehending, and reusing these mappings.

The obstacles arise from factors such as the lack of detailed documentation and metadata throughout their development lifecycle [7]. We argue that to effectively support mapping activities within communities, a design adhering to FAIR characteristics is essential [8].

As an initial step towards promoting FAIR declarative mappings and facilitating various mapping-related activities, we have introduced a comprehensive metadata model that captures the lifecycle of mapping development [9]. This metadata model is meticulously documented, versioned, and published alongside the mapping files, adhering to and upholding the FAIR data principles.

This model serves as the foundation for 'MetaSEMAP'—a tool dedicated to simplifying mapping annotation and promoting reuse. MetaSEMAP incorporates essential features, enabling users to annotate different types of mappings and comprehensively capturing the entire mapping lifecycle [10]. The ongoing aim is to enhance reuse and facilitate maintenance in subsequent phases of this research's development.

In the following section, we present a comprehensive overview of declarative mappings. Section 3 offers an examination of related work, situating MetaSEMAP within the broader context of existing research and tools in the field. Section 4 details the design and implementation of MetaSEMAP, while Section 5 discusses the usability evaluation experiment that has been conducted. Finally, Section 6 concludes the paper by summarizing key points and discussing future research directions.

2 Declarative Mappings

Declarative Mapping, within the context of this research, refers to the process of establishing connections between schemas. Serving as the foundation for interoperability and seamless data exchange, mapping involves identifying and associating corresponding elements, structures, or attributes across different information systems [1]. In the following sections, we explore mapping, commencing with the different types of mappings. Subsequently, we introduce, definition of mapping lifecycle. Additionally, we explore the concept of FAIR mappings, aligning with the principles of Findability, Accessibility, Interoperability, and Reusability, to underscore the importance of fostering a standardized and collaborative mapping ecosystem.

2.1 Declarative Mapping types

In the dynamic landscape of linked data and the semantic web, declarative mapping processes unfold in diverse patterns, typically falling into three categories: ontology mapping, interlinking, and uplift/downlift. These processes play a crucial role in establishing connections and relationships between disparate datasets, enhancing the overall interoperability and coherence of information on the web Figure 1 serves as a visual guide, providing an illustrative representation of these

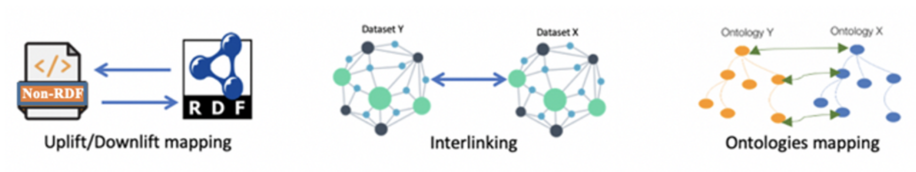

Fig. 1. The three categories of mappings.

mapping processes and their interconnections in the broader context of linked data and the semantic web.

In this study, we address all three mapping types within our proposed approach. Uplift mapping refers to converting data into RDF, with MetaSEMAP initially taking as input RDF Mapping Language (RML) mappings [3, 4]. Interlinking involves establishing connections between object descriptions to determine if they refer to the same real-world entity or have a specific relationship [2]. Ontologies mapping, or alignment, focuses on establishing correspondence between components of different ontologies, promoting semantic interoperability [11].

2.2 Mapping lifecycle

A mapping lifecycle divides the mapping process into distinct stages, and various versions of mapping lifecycles have been suggested in the literature. The lifecycle in this study draws from previously proposed lifecycles in [12], [13] and [14], but introduces a new phase, 'testing'. Consisting of six phases—Analysis, Mapping Design, Mapping Development, Mapping Testing, and Mapping Maintenance—the lifecycle is considered applicable across declarative mapping types with varying relevance based on the specific type.

2.3 FAIR declarative mappings

FAIR Data Principles were introduced in 2016 and have rapidly gained traction as a way to make digital data a more valuable [8]. Findable, Accessible, Interoperable, and Reusable are four FAIR principles, and we believe that adopting these principles when it comes to metadata for declarative mappings will help us establish a platform where semantic web and linked data community members can discover, share, maintain, and evaluate mappings. As part of the design of our proposed metadata model, the FAIR principles were taken into consideration. It begins with ensuring that the mapping and its metadata can be found both by humans and computers. This will be achieved by assigning a globally unique identifier to both the mapping and the metadata such that they can be automatically processed. Upon finding the mapping that a user needs, clear restrictions and conditions of accessibility are provided as a part of the metadata model. Then, in order to enable them to be exchanged and reused, both the mapping and metadata should have a formal, accessible, shared, and broadly accessible format and representation. Additionally, it should be clearly indicated in the mapping's associated metadata what kind of re-use is permitted. Finally, FAIR ideally aims to optimize the reuse of data and in order to accomplish this, we need a comprehensive metadata model that describes the entire development process behind these mappings. Future work on this project will focus on evaluating annotated mappings using the proposed metadata model, ensuring compliance with the FAIR principles.

3 Related Work

Several tools have been developed to support declarative mappings, particularly for uplift mappings, ontology alignment, and interlinking, but none fully address the FAIR principles for documenting these mappings.

For **uplift mappings**, R2RML and RML are widely used standards for transforming relational data into RDF [7, 15]. However, these tools lack built-in mechanisms for lifecycle documentation and metadata annotation, which are crucial for reuse and long-term sustainability. While they support data transformation, they do not fully ensure mappings are FAIR-compliant.

In **ontology alignment**, tools like AgreementMaker and OntoMerge assist in identifying correspondences between different ontologies [16, 17]. However, these tools also lack explicit support for documenting the alignment lifecycle, limiting their ability to ensure long-term reuse and compliance with FAIR principles.

For **interlinking**, tools like Silk [18] and OpenLink's Data Explorer [19] focus on linking entities across datasets. Silk supports the discovery of relationships between datasets and provides a flexible framework for interlinking data, but it lacks integrated metadata management for documenting the linking process, which limits its reusability.

Recent advancements in metadata representation, including Named Graphs and RDF-star, offer improved methods for organizing and embedding metadata [20, 21]. Named Graphs provide contextual grouping of metadata, while RDF-star enables inline annotations of triples. However, these approaches are not yet fully integrated into mapping tools, limiting their effectiveness in comprehensive metadata management.

In recent work, Trojahn (2022) explored the importance of ensuring that ontology alignments adhere to the FAIR principles, particularly in terms of metadata and documentation. The study highlights the need for alignment models that support the Findability, Accessibility, Interoperability, and Reusability of ontologies and alignments [22].

Dimou et al. (2016) introduced a system for automating metadata generation for linked data publishing workflows, which aids in ensuring the sustainability and accessibility of data through automated documentation mechanisms [23]. Similarly, Toledo et al. (2024) proposed RMLdoc, a tool designed to document mapping rules for knowledge graph construction, emphasizing the importance of documenting the mapping process to ensure its reusability and transparency in the knowledge graph construction lifecycle [24]. These efforts, while addressing metadata documentation, do not fully align with the FAIR principles for documenting declarative mappings, particularly in the context of ensuring lifecycle documentation and long-term reusability.

Another important point, given its significance in the semantic web, is the Ontology Alignment Evaluation Initiative (AOEI). While AOEI has advanced ontology alignment methods, its yearly evaluation results are published without metadata, hindering long-term discoverability and reuse. This highlights the need for proper metadata documentation to align with FAIR principles [25].

MetaSEMAP addresses these limitations by providing a tool that supports the full lifecycle of declarative mappings, including uplift mappings, ontology alignment, and interlinking, through a unified metadata model that simplifies the process, reduces ambiguity, and lowers the learning curve, while ensuring FAIR-compliant documentation for all mapping types [9].

4 Metadata Model and MetaSEMAP Tool

The solution design began with the creation of a mapping lifecycle tailored to handle three distinct mapping types detailed in Section 2.2. Building upon this lifecycle, we have developed a comprehensive metadata model to document all activities occurring throughout its phases. This metadata model serves as the cornerstone for MetaSEMAP, a tool crafted to empower knowledge engineers in annotating various mapping types through an efficient step-by-step process. The tool's future development aims to enhance additional activities, such as mapping reuse and maintenance. The subsequent section provides an in-depth exploration of the metadata model and the MetaSEMAP tool, encompassing associated design decisions.

4.1 Metadata model

To capture and document activities and decisions throughout each phase of mapping development, metadata fields were proposed [9]. Table 1 shows the structure and organization of the proposed metadata model, highlighting its alignment with the mapping lifecycle phases to ensure comprehensive documentation and facilitate reuse and maintenance.

Competency questions were formulated for each lifecycle stage to aid in validating the proposed metadata model. Following this, the model was formalized into a specification, where each field is precisely defined with an associated vocabulary term, expected value type, and a comprehensive description of its purpose. The metadata model utilizes various RDF namespaces from established ontologies. For instance, the FOAF vocabulary captured stakeholder information, including names and organizations, while the Dublin Core Metadata Initiative (DCMI) vocabulary detailed aspects such as format, source, and creator

Phases	Proposed Metadata Fields
Analysis	Metadata about stakeholders (URI, Name, Background, Role, Organization), Metadata about the purpose of the mapping (Requirements, Type of mapping, Mapping domain, Domain assumptions, Technical requirements, Risks, or Issues), and Metadata to describe inputs that will be mapped (URI, Name, Source, Type, Creator, Format).
Design	Final design decisions, Design decision justification, Quality metrics to consider during the development.
Development	URI, Name, Start/End date of the development, Tools, Mapping method, Mapping algorithm, Format.
Testing	URI, Name, Testing type, Date/Time, Testing result.
Maintenance	Publisher name, Publisher source, Version number, Version date/time.

Table 1. Proposed Metadata Fields for Each Phase

of the input file. To represent metadata related to the mapping itself, including requirements, domain, tools, and methods, a custom namespace was created.

In the first development iteration, the metadata model was described using simple vocabularies and a custom namespace. However, in the next iteration, a full ontology, using well-known reused vocabularies and ontologies, is under development. This ontology will be evaluated with knowledge engineers in the second phase of development. Vocabularies such as PROV-O [26], The Data Quality Vocabulary (DQV) [27] and the Mapping Quality Vocabulary (MQV) [28] will be used to capture provenance and metadata related to mapping quality assessment, refinement, and validation.

Additionally, a survey experiment was conducted within the semantic web and linked data community to gather feedback on the initial mapping metadata model [29]. The purpose was to evaluate the relevance of the proposed metadata fields for two mapping-related tasks. Results indicate a generally positive response. Results from the survey indicated that most of the suggested metadata fields were deemed relevant, and participants provided valuable suggestions for additional useful metadata fields.

4.2 Metadata Representation

Effective representation and attachment of metadata to RDF-expressed mappings significantly impact storage efficiency, query performance, and usability [30]. Various RDF metadata approaches, such as standard reification, named graphs, singleton properties, and RDF-star, have been explored. In [30], RDF reification approaches—standard reification, singleton property, and RDF-star—were compared based on triples, storage size, and query execution time. Results showed significant differences, with Singleton Property reducing triples by about 40%, RDF-star by almost 70%, and variations in storage size and query execution time. Additional benchmarking in [31] tested approaches on a Wikidata subset. Despite excluding Singleton Properties due to generating numerous unique predicates, RDF-star demonstrated favorable outcomes. A recent study [32] explored various knowledge graph representations, contributing insights into their impact on different consumer scenarios. Considering benchmarking contributions and recent studies, RDF-star and named graphs were chosen as representations for the proposed metadata model.

In this experiment, each mapping type—uplift mappings, ontology alignment, and interlinking—was annotated using both Named Graph and RDF-star representations. In the Named Graph approach, metadata is stored in a separate file, describing the entire mapping project, while RDF-star allows embedding metadata directly within the mapping file. For uplift mappings, each triple map was annotated separately, using a simple mapping with a single triple map for this experiment. In the second iteration, finer annotation can be applied to uplift mappings, such as adding metadata to specific triples (e.g., logical source to describe the input file). In interlinking, an RDF triple annotated the SPARQL query that links two datasets, and the same approach was applied to ontology alignment, where the RDF-star annotation captured both the alignment operation and its associated metadata in the same file. In the second iteration, MetaSEMAP will evolve to support larger and more complex mappings, and the feasibility of using RDF-star for such mappings will be assessed

4.3 MetaSEMAP

The annotation process within MetaSEMAP unfolds through four key steps. Initially, users begin the process by uploading the mapping file into the system. Next, they fill in the required metadata fields, where the system prompts users to provide relevant details. The annotation step then takes place, allowing users to view and download the annotated mappings in RDF-star or Named Graph representations.

The initial tool design and capabilities were introduced in [10], where it was first named MetaMap. The tool was primarily focused on metadata annotation for uplift mappings and aimed at supporting the reuse and lifecycle documentation of these mappings. The name was later changed to MetaSEMAP to distinguish it from other existing tools and better reflect its expanded capabilities.

MetaSEMAP integrates SPARQL 1.1 [33], enabling the querying and representation of RDF data for efficient metadata annotation. The application backend is powered by Flask, a lightweight Python web framework [34, 35], which handles metadata input and graph generation. The user interface is crafted with HTML [36] and CSS [37], delivering a clean and responsive design that simplifies the metadata annotation process for participants.

By integrating these tools and technologies, MetaSEMAP facilitates efficient and effective metadata annotation. It offers robust support for both RDF-star and Named Graph representations, promoting interoperability, clarity, and adherence to the FAIR principles.

5 Usability Evaluation Experiment

This section describes a usability evaluation performed to assess the annotation component in the MetaSEMAP tool. The following sections describe the objectives of the experiment, the design of the experiments, and how participants were presented with and guided through completing an annotation task. Next, it discusses the usability metrics and shows the evaluation results. The section ends with a summary of the key findings.

5.1 Experiment Objectives

The primary goal of this study was to evaluate the usability of MetaSEMAP, identify challenges faced, and seek opinions that will help in transforming the tool for more effectiveness in metadata annotation tasks. The objectives are as follows: Assess the usability of MetaSEMAP, Evaluate its effectiveness in metadata annotation and Identify areas for improvement.

5.2 Experiment Design

The participants (a total of 50) in the experiment were MSc. students enrolled in a course on knowledge and data engineering, with an introductory background in mapping processes. Approximately 70% of them were international students for whom English is a second language. The experiment design was approved by the university's Ethical Committee. Participants were provided with directions to carry out the experiment asynchronously and online : (MetaSEMAP Experiment). Each participant was assigned one of three randomly selected scenarios associated with the three different types of declarative mapping: uplift mapping, ontology alignment, or interlinking. Each scenario included metadata related to the Virtual Record Treasury of Ireland (VRTI), a historical knowledge graph designed to digitally reconstruct and explore Ireland's archival heritage. The VRTI knowledge graph [38] was chosen as the basis for the scenarios because it is openly available, easily understandable and accessible to master's-level students, even for those with a more technical background rather than a humanities background. This makes it particularly suitable for exploring metadata annotation processes in mapping scenarios due to its rich historical context, structured representation of archival data, and relevance for declarative mapping tasks.

Participants started by downloading the mapping file corresponding to their assigned scenario and reading the scenario description to understand the context of the mapping project. They then uploaded the mapping file to the MetaSEMAP tool, examined its content, and populated the metadata fields using the scenario as a reference, Figure 2 presents an example of the interface, where users input metadata.

For space reason, The three scenarios provided to participants are summarized in Table 2. Each scenario corresponded to a specific mapping type—uplift mapping, ontology alignment, or interlinking—and focused on enhancing the Virtual Record Treasury of Ireland (VRTI) Knowledge Graph.

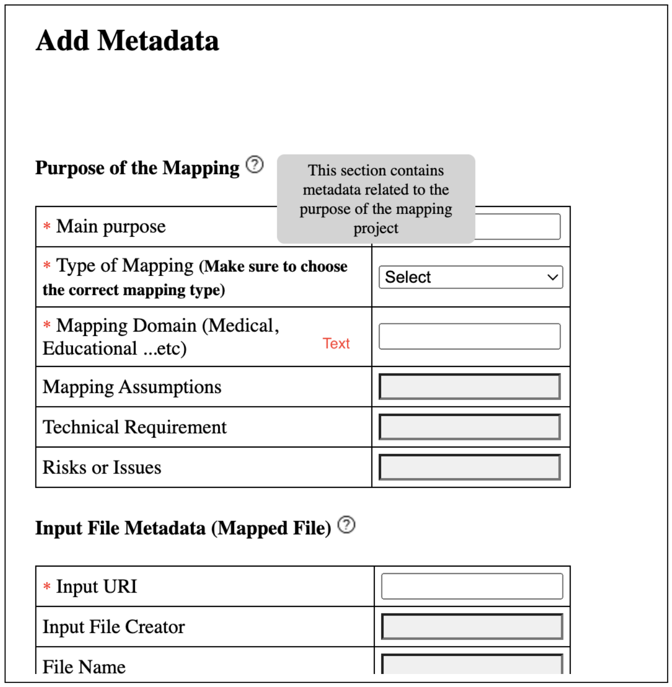

Fig. 2. MetaSEMAP Interface for Metadata Annotation.

Table 2. Summary of Scenarios in the Usability Experiment

Mapping Type	Summary of Scenario
Uplift Mapping	Convert county information from historical data into RDF triples. The data file `data_county.csv` was used to map county IDs, names, and geographic details to enrich the VRTI Knowledge Graph.
Ontology Alignment	Align person entities (e.g., historical figures) between the VRTI Knowledge Graph and external datasets. Features such as relationships, affiliations, and roles were matched using the file `person_alignment.rdf`.
Interlinking	Connect Irish historical figures to their corresponding Wikidata entries using owl:sameAs relationships. The project enriched VRTI records with biographical and contextual data using the file `interlinking.rdf`.

After populating the metadata, participants verified their inputs and submitted metadata entries. The results of the annotation were generated in two representation: RDF-star and Named Graph. Participants could view or download the automatically annotated file in their preferred representation . Listing 1.1 and Listing 1.2 show part of the metadata fields populated by different participants in both representations for uplift mapping:

```
graph:uplift_mapping {
    metag:subject metag:endDate "2024-02-28"^^xsd:date ;
        metag:fileFormat "CSV" ;
        metag:finalDesignDecisions "create new mapping" ;
        metag:mappingDomain "Archival" ;
        metag:mappingMethod "Automatic" ;
        metag:mappingType "Uplift Mapping" ;
        metag:mappingURI "http://virtualrecordtreasury.ie/uplift_mapping_v1.0" ;
        metag:publisherName "VRTI team" ;
        metag:publisherSource "https://virtualtreasury.ie" ;
        metag:purpose "Uplift Mapping for Irish Counties for VRTI" ;
        metag:qualityMetrics "Well outlined geospatial and semantic information for Irish counties" ;
        metag:startDate "2024-01-01"^^xsd:date ;
        metag:testingDate "2024-02-28"^^xsd:date ;
        metag:testingResult "Successful" ;
        metag:testingType "Validation" ;
        metag:versionDateTime "2024-02-28"^^xsd:date ;
        metag:versionNumber "1.0" ;
        dcmi:source "http://virtualrecordtreasury.ie/" .}
```

Listing 1.1. Named Graph Metadata for Uplift Mapping

```
# Metadata annotations for the uplift mapping operation
<< <http://example.com/ns##COUNTY> <http://www.w3.org/1999/02/22-rdf-syntax-ns#type> <http://www.w3.org/ns/r2rml#TriplesMap> >>
    <http://example.com/metag/purpose> "Uplift Mapping for Irish Counties in the Virtual Record Treasury of Ireland" ;
    <http://example.com/metag/mappingType> "Uplift Mapping" ;
    <http://example.com/metag/mappingDomain> "historical" ;
    <http://purl.org/dc/terms/source> "data_county.csv" ;
    <http://example.com/metag/mappingMethod> "Automatic" ;
    <http://example.com/metag/mappingURI> "http://virtualrecordtreasury.ie/uplift_mapping_v1.0" ;
    <http://example.com/metag/testingType> "Validation" ;
    <http://example.com/metag/testingDate> "2024-02-28" ;
    <http://example.com/metag/testingResult> "Successful" ;
    <http://example.com/metag/publisherName> "VRTI team" ;
    <http://example.com/metag/publisherSource> " https://virtualtreasury.ie" ;
    <http://example.com/metag/versionNumber> "1.0" ;
    <http://example.com/metag/versionDateTime> "2024-02-28" .
```

Listing 1.2. RDF-star Metadata for Uplift Mapping

Finally, participants completed a usability survey to evaluate the annotation component, provide feedback for improvements, and indicate their preferred metadata representation.

5.3 Experiment Metrics

The metrics used for the evaluation of this usability experiment included both quantitative and qualitative measures. The Post Study Usability Questionnaire

(PSSUQ) [39] was used to assess user satisfaction with the MetaSEMAP tool. The PSSUQ consists of 14 Likert scale questions, evaluating aspects such as system usability, ease of learning, error recovery, and overall satisfaction. Additionally, participants were asked to indicate which metadata representation (RDF-star or Named Graph) they found easier to read and understand. An open comment section allowed participants to provide additional feedback on the tool's design and functionality.

Further metrics included task completion time, which measured the duration taken to annotate mappings and complete all necessary metadata fields, and error rate, which tracked the frequency and types of issues encountered during the experiment. Qualitative data from open-ended responses were analyzed using thematic analysis to identify recurring patterns, user perceptions, and suggestions for improving the tool. Table 3 summarizes these metrics and how they were measured.

Table 3. Summary of Evaluation Metrics

Metric	What is Being Measured	How It Was Measured
System Usability, Ease of Learning, Error Recovery, Overall Satisfaction	Usability, ease of use, learning, error recovery, and overall satisfaction with the tool	Measured through the PSSUQ, which includes 14 Likert scale questions assessing various aspects of tool usability.
Task Completion Time	Time taken to annotate mappings and complete metadata fields	Recorded by tracking the duration of the task from start to finish.
Error Rate	Frequency and type of issues encountered during the experiment	Counted through errors encountered.
Preferred Metadata Representation	Participant preference between RDF-star and Named Graph formats	Measured by asking participants which format they found easier to read and understand during the annotation task.
Thematic Analysis	Insights and suggestions from user feedback	Analyzed open-ended responses to identify common themes and areas for improvement.

5.4 Results and Discussion

System Usability and User Satisfaction The overall system usability was assessed using the Post-Study System Usability Questionnaire (PSSUQ), which was represented using a 1 to 7 Likert scale. A total of 46 out of 50 students successfully completed the survey after the experiment, with an average score of 2.6 (where lower scores indicate better usability). This suggests that participants generally found the tool effective and user-friendly. The survey revealed that approximately 70% of participants rated the tool positively for ease of use, particularly appreciating its simplicity for annotating mapping files.

The System Usefulness category (average score: 2.4) indicates that MetaSEMAP was deemed useful, with participants able to recover from errors quickly and feeling confident while annotating mapping files. The scores for Information Quality (average score: 3.4) reflect moderate satisfaction, highlighting areas for improvement, such as the clarity of of the error messages. The Interface Quality category received a positive response, with an average score of 2.3, indicating that the interface was intuitive, consistent, and easy to navigate. Table 4 summarizes these average scores, providing a clear overview of the tool's usability performance.

These results suggest that future iterations of MetaSEMAP should focus on improving the clarity of error messages and reducing perceived complexity while maintaining the system's overall ease of use.

Category	**Average Score**
System Usefulness	2.4
Information Quality	3.4
Interface Quality	2.3
Overall Usability	2.6

Table 4. Average Scores for Each PSSUQ Category

Task Completion Time A total of 50 participants were involved in the experiment. On average, participants took approximately 44 minutes to complete the annotation task, which started after reading the scenario of the mapping project and the experiment instructions. The task began with filling in the appropriate metadata fields and ended once participants reviewed and submitted the filled metadata fields.

The time varied depending on the complexity of the task. Participants working on Ontology Alignment mappings spent more time (40 minutes on average) compared to those working on Uplift Mapping (28 minutes on average) and Interlinking mappings (33 minutes on average).

Extreme outliers were identified in the data. Some participants completed the task in unusually low times (e.g., under 5 minutes, such as 89 seconds, 193 seconds, and 77 seconds), likely indicating incomplete or rushed tasks. Conversely, some participants took unusually high times (e.g., over 200 minutes, such as 12291 seconds, 29251 seconds, and 29291 seconds), likely due to task-related difficulties or errors. These extreme times were excluded from the analysis to provide a more accurate reflection of the average task completion time. After excluding these outliers, the revised average completion time was approximately

32 minutes for Uplift Mapping, 39 minutes for Ontology Alignment, and 33 minutes for Interlinking.

Error Rate The error rate in this context refers to instances where participants either misclassified metadata fields, provided incomplete or incorrect values, or omitted required information. Participants were tasked with completing key metadata fields for three scenarios: *Uplift Mapping*, *Ontology Alignment*, and *Interlinking*. The evaluation assessed their submissions for correctness by comparing inputs against predefined values for mandatory fields. Results were categorized into *Correct*, *Partially Correct*, and *Incorrect*. Percentages were calculated as follows:

$$Correct\% = \left(\frac{NumberofCorrectResponses}{TotalFields}\right) \times 100 \quad (1)$$

The table below summarizes the results for each scenario:

Table 5. Participants' Performance Across Scenarios

Scenario	**Correct (%)**	**Partially Correct (%)**	**Incorrect (%)**
Uplift Mapping	65	25	10
Ontology Alignment	60	28	12
Interlinking	68	20	12

The results demonstrate that participants exhibited strong performance in completing key metadata fields. Correct answers were consistently high across all three scenarios, with the *Interlinking* scenario yielding the highest accuracy (68%). However, *Partially Correct* responses highlight areas where participants understood the intent but struggled with precision. For instance, in the *Uplift Mapping* scenario, fields such as *purpose* and *qualityMetrics* often featured minor deviations in phrasing but still conveyed the intended meaning. Similarly, in the *Ontology Alignment* scenario, broader terms like "Historical Records" were accepted for the *mappingDomain* field, although more precise terminology was expected.

Incorrect fields reflect specific challenges. This was particularly notable in the *Interlinking* scenario, where the distinction between *manual* and *automatic* mapping methods caused confusion. Additionally, fields requiring descriptive inputs, such as *qualityMetrics* and *purpose*, tended to include overly generic or simplified responses.

In general, participants demonstrated a solid understanding of fundamental metadata fields, particularly in identifying *mapping purpose*, *dates*, and *publisher names*. Structured fields with clear formats, such as *dates* and *URIs*, had high correctness rates, reflecting their clarity. Free-text fields like *purpose* and *qualityMetrics* showed variations in responses, highlighting the need for standardized templates or clearer examples.

While participants successfully completed core metadata tasks, the results suggest that the tool could benefit from further enhancements to address specific challenges. Providing in-tool *guidance*, *examples*, or *pre-filled templates* for complex fields can reduce errors and improve precision. Additionally, clearer definitions and distinctions (e.g., *manual* vs. *automatic mapping methods*) would help mitigate confusion. These insights underscore the tool's effectiveness for basic annotation tasks while identifying areas for refinement to improve usability for more technical fields.

Preferred Metadata Representation The survey results revealed a strong preference for Named Graph as the metadata representation. 74.4% of participants preferred Named Graph for its ease of understanding. Named Graph offers a more clear structure independent from the mapping statements, which participants found helpful in organizing and interpreting the metadata. This preference aligns with the tool's design, which emphasizes clarity and structure, especially for users familiar with graph-based representations.

In contrast, 25.6% of participants preferred RDF-star, which was appreciated for its simplified structure, and compact syntax. RDF-star allows metadata to be embedded directly within triples, providing immediate context and reducing the need for external references or complex graph structures, which made it easier for some users to follow annotations.

Keyword-Based Thematic Analysis By extracting keywords from participants' feedback, we identified several recurring themes and areas for improvement:

- **Ease of Use / Simplicity:**
  - *Keywords:* easy to use, non-technical, simple, intuitive
  - *Insights:* MetaSEMAP was generally seen as user-friendly and accessible to non-technical users.
- **Error Handling:**
  - *Keywords:* error messages, TypeError, error recovery, cleaner error handling, validator
  - *Insights:* Users encountered errors and suggested improved error messages, clearer error recovery mechanisms, and validators for fields like URIs.
- **Field Guidance:**
  - *Keywords:* tooltips, field descriptions, explanations, help
  - *Insights:* Participants recommended adding more tooltips or explanations for metadata fields to help users understand what is required.
- **Interface:**
  - *Keywords:* interface, size of input boxes, text wrapping, scenario visibility, design
  - *Insights:* Several users recommended improvements to the interface, such as larger input boxes, text wrapping, and visibility of the scenario alongside the form.

- **Additional Features:**
  - *Keywords:* templates, collaboration, version control, visualization tools, consistency checks
  - *Insights:* Users suggested adding features like predefined templates, better collaboration tools, search/filter options, and enhanced consistency checks.

This thematic analysis highlights the key areas for improving MetaSEMAP, with a particular focus on error handling, field guidance, and interface enhancements. These insights can guide future iterations of the tool to better meet user needs and improve overall usability.

6 Conclusion and Future Work

This paper presented the usability evaluation of MetaSEMAP, a tool designed to facilitate the annotation of mapping projects and enhance their reusability according to the FAIR principles. The evaluation focused on assessing the tool's usability using the Post-Study System Usability Questionnaire (PSSUQ), with results showing high user satisfaction. Participants found the tool intuitive and effective, particularly appreciating its simplicity for annotating mapping files.

The analysis highlighted key strengths in the tool's usability, including ease of use, system usefulness, and interface clarity. However, challenges were noted in error recovery, which was identified as an area for improvement. Participants recommended clearer error messages, enhanced error recovery mechanisms, and validators for required fields like URIs.

The task completion time varied depending on the complexity of the mapping types, with Ontology Alignment taking the most time, while Uplift Mappings were completed more quickly. This suggests that the tool's design should continue to cater to different mapping types to optimize user performance across various tasks.

Future work will focus on addressing the feedback received during this evaluation. This includes improving error handling, providing more field guidance, refining the interface, and adding additional features such as collaboration tools, version control, and validation mechanisms. Further user studies will be conducted to explore the tool's effectiveness with a broader range of real-world use cases.

In addition to metadata annotation, the annotated mappings will feed into the tool's reuse component. In this phase, users can search by filling in metadata fields or writing natural language queries, which will be converted into SPARQL queries. This component will be evaluated to assess how effectively the metadata model supports reuse decisions.

In summary, MetaSEMAP has demonstrated strong usability and effectiveness in supporting metadata annotation for declarative mappings, and the insights from this study will guide future development to enhance its capabilities and user experience.

Acknowledgments. The first author acknowledges financial support from Al Imam Mohammad Ibn Saud Islamic University and the Saudi government - represented by the Royal Embassy of Saudi Arabia - Cultural Bureau in Dublin. The second author is partially supported by the SFI ADAPT Research Centre (grant number 13/RC/2106$_P$2).

References

1. Z. Bellahsene, A. Bonifati, F. Duchateau, and Y. Velegrakis, "On evaluating schema matching and mapping," in *Schema Matching and Mapping*, Z. Bellahsene, A. Bonifati, and E. Rahm, Eds. Springer Berlin Heidelberg, 2011, pp. 253–291.
2. A. Ferrara, A. Nikolov, and F. Scharffe, "Data linking for the semantic web," *International Journal on Semantic Web and Information Systems (IJSWIS)*, vol. 7, no. 3, pp. 46–76, 2011.
3. A. Junior, C. Debruyne, R. Brennan, and D. O'Sullivan, "An evaluation of uplift mapping languages," *International Journal of Web Information Systems*, vol. 13, pp. 00–00, 2017.
4. "Rdf mapping language (rml)," Online, 2023, accessed: Mar. 29, 2023. [Online]. Available: https://rml.io/specs/rml/
5. D. Lewis, J. Keeney, D. O'Sullivan, and S. Guo, "Towards a managed extensible control plane for knowledge-based networking," in *17th IFIP/IEEE International Workshop on Distributed Systems: Operations and Management (DSOM 2006)*, 2006.
6. J. Wu, F. Orlandi, D. O'Sullivan, and S. Dev, "An ontology model for climatic data analysis," in *IEEE International Geoscience and Remote Sensing Symposium IGARSS*, 2021.
7. A. Dimou, M. Vander Sande, P. Colpaert, R. Verborgh, E. Mannens, and R. Van de Walle, "Rml: A generic language for integrated rdf mappings of heterogeneous data," in *Proceedings of the Workshop on Linked Data on the Web (LDOW)*, 2014.
8. M. D. Wilkinson *et al.*, "The fair guiding principles for scientific data management and stewardship," *Scientific data*, vol. 3, no. 1, pp. 1–9, 2016.
9. S. Alzahrani and D. O'Sullivan, "Towards a unified metadata model for semantic and data mappings," in *Ontology Matching workshop at ISWC 2022*, 2022.
10. S. Alzahrani and D. O'Sullivan, "Metamap: Facilitating mapping reuse and lifecycle documentation," in *2024 IEEE 18th International Conference on Semantic Computing (ICSC)*, 2024, pp. 323–328.
11. J. Euzenat, "An api for ontology alignment," in *The Semantic Web – ISWC 2004*, S. A. McIlraith, D. Plexousakis, and F. van Harmelen, Eds. Springer Berlin Heidelberg, 2004, pp. 698–712.
12. D. O'Sullivan, V. Wade, and D. Lewis, "Understanding as we roam," *Internet Computing, IEEE*, vol. 11, pp. 26–33, 2007.
13. H. Thomas, R. Brennan, and D. O'Sullivan, "Moom – a prototype framework for management of ontology mappings," in *Proceedings - International Conference on Advanced Information Networking and Applications, AINA*, 2011, pp. 548–555.
14. C. Debruyne, B. Walshe, and D. O'Sullivan, "Towards a project centric metadata model and lifecycle for ontology mapping governance," in *Proceedings of the 17th International Conference on Information Integration and Web-based Applications & Services (iiWAS '15)*. Association for Computing Machinery, 2015, p. Article 50.
15. W3C, "R2rml: Rdb to rdf mapping language," W3C Recommendation, 2012. [Online]. Available: https://www.w3.org/TR/r2rml/

16. J. Yang *et al.*, “Agreementmaker: A framework for ontology matching,” in *Proceedings of the International Conference on Data Engineering*. IEEE, 2011, pp. 673–683.
17. P. Shvaiko and J. Euzenat, “Ontology matching: State of the art and future challenges,” in *Proceedings of the 10th International Conference on Database Theory*, 2008, pp. 3–18.
18. R. Volz *et al.*, “Silk: A link discovery framework for the web of data,” in *Proceedings of the International Semantic Web Conference*, 2009.
19. O. Software, “Openlink data explorer,” https://www.openlinksw.com/products/data-explorer/, n.d., accessed: 2023-03-18.
20. J. J. Carroll, C. Bizer, P. Hayes, and P. Stickler, “Named graphs, provenance and trust,” in *Proceedings of the 14th International Conference on World Wide Web (WWW)*. ACM, 2005, pp. 613–622.
21. O. Hartig, “Reconciliation of rdf* and property graphs,” *arXiv preprint arXiv:1409.3148*, 2017.
22. C. Trojahn, “Fair ontologies, fair ontology alignments,” *EKAW Companion*, 2022.
23. A. Dimou, T. De Nies, R. Verborgh, E. Mannens, P. Mechant, and R. Van de Walle, “Automated metadata generation for linked data generation and publishing workflows,” vol. 1593, 2016.
24. J. Toledo, A. Iglesias-Molina, D. Chaves-Fraga, and D. Garijo, “Rmldoc: Documenting mapping rules for knowledge graph construction,” in *ESWC24 Posters and Demos Track*, 2024.
25. AOEI, “Ontology alignment evaluation initiative (aoei),” https://www.cs.ox.ac.uk/activities/ontology-alignment-evaluation/, 2023, accessed: 2023-03-18.
26. L. Moreau *et al.*, “Prov-o: The provenance ontology,” 2013. [Online]. Available: https://www.w3.org/TR/prov-o/
27. C. L. F. Rocha *et al.*, “Data quality vocabulary (dqv),” 2015. [Online]. Available: https://www.w3.org/TR/vocab-dqv/
28. M. Randles *et al.*, “Mapping quality vocabulary (mqv): A domain-specific model for capturing metadata and provenance relating to mapping quality assessment, refinement, and validation,” *Journal of Semantic Web*, vol. 15, no. 4, pp. 305–323, 2021. [Online]. Available: https://doi.org/10.1007/s11220-021-00325-4
29. S. Alzahrani and D. O’Sullivan, “What metadata is needed for semantic and data mappings?” in *2023 IEEE 17th International Conference on Semantic Computing (ICSC)*, 2023, pp. 312–316.
30. F. Orlandi, D. Graux, and D. O’Sullivan, “Benchmarking rdf metadata representations: Reification, singleton property and rdf,” in *2021 IEEE 15th International Conference on Semantic Computing (ICSC)*. IEEE, 2021, pp. 233–240.
31. “What is rdf-star?” Online, 2023, accessed: Apr. 26, 2023. [Online]. Available: https://www.ontotext.com/knowledgehub/fundamentals/what-is-rdf-star/
32. A. Iglesias-Molina, K. Ahrabian, F. Ilievski, J. Pujara, and C. O, “Comparison of knowledge graph representations for consumer scenarios,” in *International Semantic Web Conference*. Springer, 2023, pp. 271–289.
33. W3C, “Sparql 1.1 query language,” W3C Recommendation, 2013. [Online]. Available: https://www.w3.org/TR/sparql11-query/
34. Flask Community, “Flask web framework documentation,” Online, 2023. [Online]. Available: https://flask.palletsprojects.com/
35. M. Grinberg, *Flask Web Development: Developing Web Applications with Python*, 2nd ed. O’Reilly Media, 2018.
36. W3C Community, “Html5 specification,” 2023. [Online]. Available: https://html.spec.whatwg.org/

37. Mozilla Developer Network (MDN), "Css: Cascading style sheets documentation," 2023. [Online]. Available: https://developer.mozilla.org/en-US/docs/Web/CSS
38. A. Randles, L. McKenna, L. Kilgallon, B. Yaman, P. Crooks, and D. O'Sullivan, "The knowledge graph explorer for the virtual record treasury of ireland," in *VOILA Workshop - International Semantic Web Conference (ISWC)*. CEUR, 2024, pp. 1–15.
39. J. R. Lewis, "Ibm computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use," *International Journal of Human-Computer Interaction*, vol. 7, no. 1, pp. 57–78, 1995.