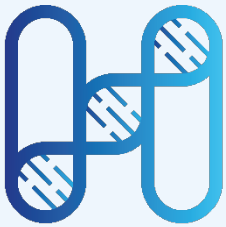

HEREDITARY

HetERogeneous sEmantic Data integratIon for the guT-bRain interplaY

Deliverable 5.7

Requirement Analysis and User Studies: Initial Results

Version 1.10, 23.12.2025

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No GA 101137074. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

Funded by the European Union

EXECUTIVE SUMMARY

In HEREDITARY's Work Package 5 (WP5) we develop and evaluate data visualization and information extraction methods, to support analysis and understanding of multimodal biomedical data. *Requirement Analysis* is a process focused on obtaining, from domain experts, requirements that data analysis and visualization approaches should fulfill. These guide the development, allow to measure the success, and evaluate the results. Requirement analysis is a continuous process.

In the Deliverables 5.1 and 5.3 of Month 12, based on preliminary requirement analysis with project partners, visualization components covering a wide range of data modalities were developed. In the course of HetERogeneous sEmantic Data integratIon for the guT-bRain interplaY (HEREDITARY), these components will gradually be developed into data visualization applications, for application and evaluation. In this deliverable we describe our activities and results in requirement engineering, to inform the application development. The activities to date led to varying levels of completeness of the visualization components and prototypes. Some components are already quite mature (e.g., the gut-brain visualization component, see Deliverable 5.4), while other have begun later and are on a conceptual and pre-prototypical implementation level yet (e.g., Section 5 in this document).

If a certain degree of maturity has been achieved for a component, user studies can be performed, assessing in a quantitative and/or qualitative way how requirements are met, how the approach performs in general, and to identify further development potential.

This deliverable summarizes work in general initial requirement analysis and user studies from a workshop conducted in March 2025. Also, a user study and requirement analysis for exploring knowledge graphs are presented, as well as a requirement analysis for data analysis in cell-cell communication patterns. Furthermore, a concept for a user study to assess and develop the Droplets image visualization technique is presented. The requirement analysis and evaluation efforts provide that development is aligned both with the state of the art, and with application requirements by analysts and domain scientists.

DOCUMENT INFORMATION

Deliverable ID	5.7
Deliverable Title	Requirement Analysis and User Studies: Initial Results
Work Package	WP 5
Lead Partner	TU Graz
Due Date	31.12.2025
Date of submission	23.12.2025
Deliverable Type	R – Report
Dissemination level	PU – Public

AUTHORS

Name	Organization
Tobias Schreck	TUGRAZ
Stefan Lengauer	TUGRAZ
Peter Waldert	TUGRAZ
Benedikt Kantz	TUGRAZ
Gaia Tussardi	UNIPD
Aleksis Datseris	ONTO
Svetla Boytcheva	ONTO
Ivelina Nikolova-Koleva	ONTO
Daniele Dell'Aglio (internal reviewer)	AAU
Anna Romanovych (contributor)	UNIPD
Gianmaria Silvello (contributor)	UNIPD

REVISION HISTORY

Version	Date	Author	Description
0.10	2025-02-28	Peter Waldert	Structure for the deliverable
0.20	2025-11-24	All Authors	Complete draft including all components
0.30	2025-12-11	All Authors	Revised version based on reviewer comments
0.35	2025-12-12	Tobias Schreck	Edited and added to executive summary, introduction, and conclusions
0.40	2025-12-16	Peter Waldert	Further corrections, extend acronyms, consistent writing style & referencing
1.00	2025-12-17	All Authors	Final version
1.10	2025-12-17	Gianmaria Silvello	Revision and submitted version

Views and opinions expressed are however those of the author(s) only and do not necessarily re ect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

Contents

1 Introduction 9

2 Survey of Application Requirements from Research Partners 9
2.1 Survey Results . . . 10
2.2 Survey Summary . . . 11

3 OnSET User Study: Interactive Querying in Knowledge Graphs 11
3.1 Related Works . . . 12
3.2 De nitions . . . 12
3.3 User Interface . . . 13
3.3.1 NLI . . . 13
3.3.2 Visual Query Builder . . . 13
3.4 Study Design . . . 14
3.4.1 User Tasks . . . 14
3.5 Study Results . . . 16
3.6 Discussion . . . 16

4 Talk to Your Graph: Natural Language Querying of Neurodegenerative Diseases 17
4.1 Methodology . . . 17
4.2 Literature Review . . . 17
4.3 Focus Group . . . 18
4.4 Results . . . 18
4.4.1 Must . . . 18
4.4.2 Nice-to-Have . . . 19
4.5 Discussion and Future Work . . . 19

5 Concept for Visual Analysis of Cell-Cell Signaling 21
5.1 Multi-layer Network Visualization . . . 22
5.2 Dashboard Design . . . 22
5.3 Next Steps . . . 24

6 Initial Design of a User Study on Glyph Design 24
6.1 Related Work . . . 25
6.2 Methodology . . . 26
6.3 Next Steps . . . 27

7 Conclusion 27

References 28

List of Figures

1 Distribution of Responses for HEREDITARY Interests (proportional to the number of participants who voted for these multiple-choice questions). . . . 10

2 Distribution of Responses for Data Modalities of Interest (proportional to the number of participants who voted for these multiple-choice questions). 10
3 Distribution of Responses for Existing Visualisations of Interest (proportional to the number of participants who voted for these multiple-choice questions). 11
4 A simple example of a prototype graph $G_p := (N_p; E_p)$ with three nodes and two links. 13
5 Prototype Graph extraction process from query. 14
6 Ontology and Sparse data Exploration Tool (OnSET) user interface (Kantz et al., 2025b) showing a query. 15
7 Figure of the Talk To Your Graphs (TTYGs) UI 19
8 Graphical illustration of (a) cell-cell signaling conceptualization and modeling through a multi-layer network, (b) dashboard design, and (c) schematic example of biological hypothesis drawn from multi-layer network interpretation. Created with BioRender.com 23
9 Screenshot of a first draft of the dashboard, showing an initial implementation of the concentric circular layout. 23
10 Droplets glyph design (a) showing spatial transcriptomics data of a coronal cut of a mouse brain, superimposed to a Hematoxylin and Eosin stain (b). A side-by-side comparison of a baseline pie chart visualization (c) with our solution (d) underlines the benefit of the latter. . . . 25

Acronyms

2AFC Two-Alternative Forced Choice. 26

ALS Amyotrophic Lateral Sclerosis. 11

BGP Basic Graph Pattern. 12

Droplet Dense Representation of Proportion Landscapes Enhancing Trends. 24–26

EEG Electroencephalography. 11

FDG-PET Fluorodeoxyglucose Positron Emission Tomography. 11

HEREDITARY HetERogeneous sEmantic Data integratIon for the guT-bRain interplaY. 3, 9, 11, 12, 17, 18, 20, 27

KG Knowledge Graph. 11–14, 16–18

LLDI Linked Life Data Inventory. 17

LLM Large Language Model. 12–14, 17–20, 22, 24

MRI Magnetic Resonance Imaging. 11

NLI Natural Language Interface. 12–14

NLQ Natural Language Querying. 17, 19

OnSET Ontology and Sparse data Exploration Tool. 7, 9, 12–16

SPARQL SPARQL Protocol and RDF Query Language. 12, 17, 19, 21

SUS System Usability Score. 12, 14, 16

TTYG Talk To Your Graph. 7, 17–19

VQ Visual Querying. 17

WGS Whole Genome Sequencing. 11

WP Work Package. 11, 12

WP5 Work Package 5. 3, 9

1 Introduction

In HEREDITARY's WP5 we develop and evaluate data visualization and information extraction methods, to support analysis and understanding of multimodal biomedical data. In the Deliverables 5.1 and 5.3 of month 12, based on initial requirement analysis with project partners, visualization components covering a wide range of data modalities were developed. In the course of HEREDITARY, these components will gradually be developed into data visualization applications, for application and evaluation. In this deliverable we describe our activities and results in requirement engineering, to inform the application development.

Requirement analysis is a process to obtain, from domain experts, requirements and expectations put forward to a solution which is not yet achieved, but which is desired. That solution should ll a gap, improve over existing solutions, or enable new activities. Designing components for visual analysis of multi modal biomedical data, our approaches developed in WP5 depend on requirement analysis to inform them. As we will gradually develop and re ne our visualization components into standalone applications, the requirement analysis process is done continuously. Depending on the component, and when it was started, the requirements take on different degrees of completeness and detail. In some cases they are quite robust, in others these are a beginning. In this sense, this deliverable has results on requirement analysis on different states.

Early in the project, we conducted a project workshop to showcase early prototypes of various visualization components, and we asked the participants (all from research partners) for their interests and requirements on a high-level. In Section 2 the results of this survey are presented. These are high-level requirements that inform the WP5 research.

A signi cant part of work in WP5 addresses searching and exploring in knowledge graph data extracted from scienti c literature. These include the OnSET prototype, and works by partner OntoText. The respective components are already quite mature and can be employed by parners for search and exploration. Hence, we could perform user studies on these components. Sections 3 and 4 present respective evaluation and requirement analysis.

A rather new line of development considers network visualization of speci c types of cell-to-cell interaction patterns. In Section 5, we present results of an initial requirement analysis and visualization design for a new visualization component which can be applied in this area. Speci ally, we present a network visualization making use of hierarchical network visualization, in conjunction with advanced user interaction approaches.

Already in year one, a novel image visualization technique (`Droplets') was proposed, see Deliverable 5.3. As a next step, we want to perform a user study, assessing the effectiveness of the Droplets visualization in exploring image-based data. To this end, in Section 6, we develop a concept for a user study, aiming at validating the original Droplet visualization design and informing future improvements.

Section 7 concludes and summarizes next steps.

2 Survey of Application Requirements from Research Partners

The visualization components that were developed during Y1 in HEREDITARY were informed by the working program and research plan, and by bilateral initial requirement analysis with several partners (see Deliverables 5.1 and 5.3). To develop the requirement analysis, in March 20, 2025, we conducted an online Half-Day workshop for the HEREDITARY consortium, presenting the visualization components developed during Y1 (also see Deliverables 5.1 and 5.3). About 30 participants attended the workshop, representing HEREDITARY partners and researchers. The goal was to inform about the components and obtain a feedback on application possibilities.

In preparation of the workshop, a survey was conducted inquiring about application requirements of visualization for data analysis in the project. Our survey is loosely based on the design triangle analysis method

(Miksch & Aigner, 2014). This method suggests to describe requirements in terms of data, users, and tasks. We received 26 responses in total, giving an idea about requirements, and informing follow-up requirement meetings.

2.1 Survey Results

The survey was structured into questions regarding applications and visualization interest, and into available to be analyzed.

Interests of Partners We asked participants (1) for their main application interests, (2) data modalities of interest, and (3) visualization components of interest. The percentage-aggregated results for the multiple-choice questionnaire on individual participants' interests can be found in Figures 1 to 3.

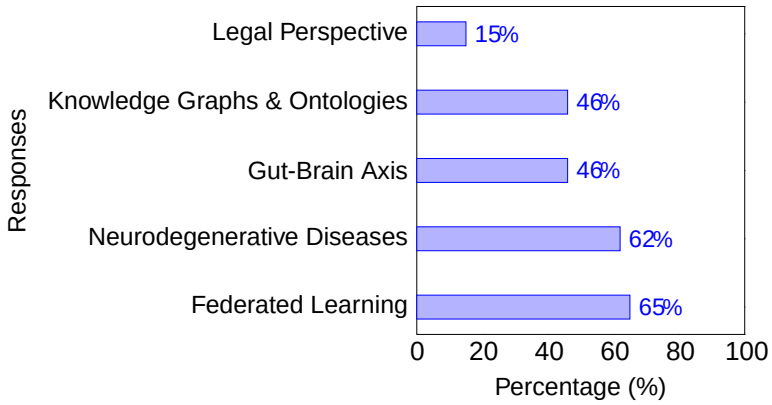

Figure 1: Distribution of Responses for HEREDITARY Interests (proportional to the number of participants who voted for these multiple-choice questions).

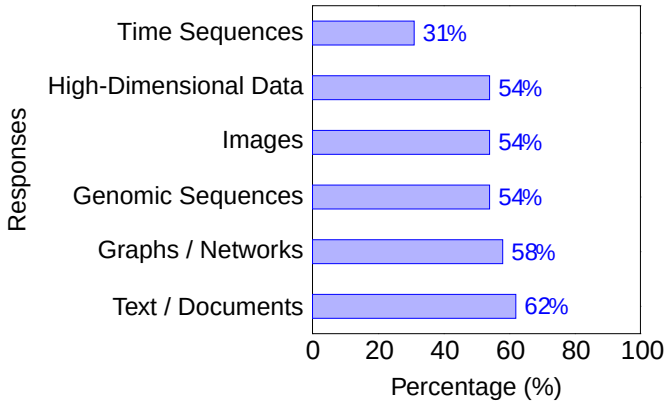

Figure 2: Distribution of Responses for Data Modalities of Interest (proportional to the number of participants who voted for these multiple-choice questions).

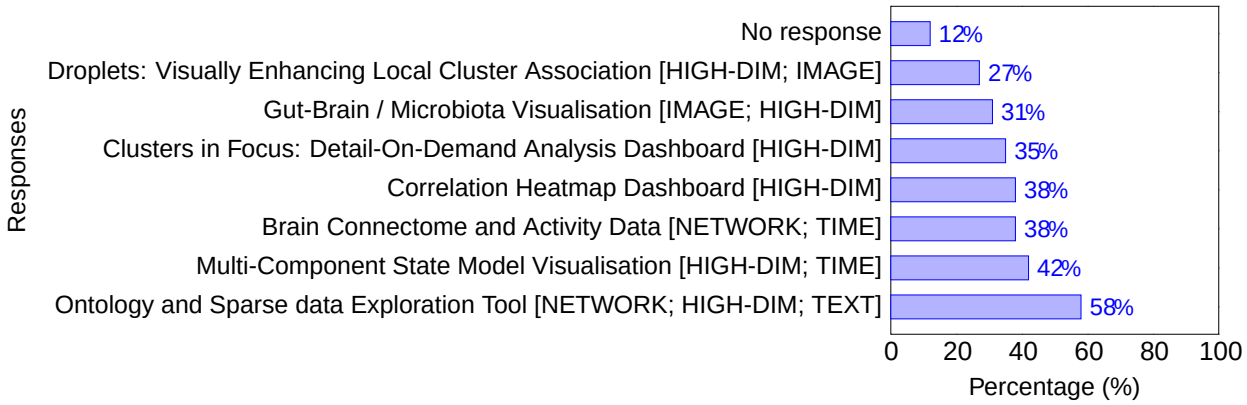

Figure 3: Distribution of Responses for Existing Visualisations of Interest (proportional to the number of participants who voted for these multiple-choice questions).

Data Availabilities of Project Partners This section of the questionaire collected information about potentially available data types and modalities from participants, for use in visualization technique design and evaluation.

- Multimodal ALS clinical data (clinical, biomarker, respiratory measures), genetic (WGS), neuroimaging (FDG-PET, MRI)
- Brain imaging, connectomics, microbiota
- AnswerALS data (genomic, transcriptomic, epigenomic, proteomic)
- Neurodegenerative Multi-omics
- Single-cell communication data linked with transcription factors activity
- EEG data; deep learning architectures; embedding space

2.2 Survey Summary

The conducted survey indicates a high interest in visualisation as part of the HEREDITARY project, mostly concentrated around the central themes of the project: Federated Learning, Neurodegenerative Diseases and the Gut-Brain Axis. Data Modality types indicate which types of visualisations to further develop / focus on in the future, mostly concentrated around textual data, but also graphs. The responses also indicate a generally high interest in the existing visualisation tools and their usefulness.

This survey con rmed the initially taken directions in developing visualization components. The workshop and survey fostered the understanding of the potential of data visualization in HEREDITARY to solve analysis problems. Many research directions were discussed in the workshop which inform the future requirement analysis and evaluation efforts.

3 OnSET User Study: Interactive Querying in Knowledge Graphs

Semantic exploration systems for ontologies and Knowledge Graphs (KGs) are an integral part of the HEREDITARY Work Packages (WPs), as these provide gateways for users to interact and query these knowledge

bases. OnSET serves as one of the core querying components used within HEREDITARY to enable users to query the KGs created by WP 4 (Nikolova-Koleva et al., 2025; Schreck, Lengauer, Waldert, Kantz, Boytcheva, et al., 2025). OnSET allows users with varying degree of knowledge of KGs, querying interfaces, and expertise in the application domain to query databases. Our prior work has, speci cally, explored the possibility to create queries using Natural Language Interface (NLI) to enable users to enter a base query formulating their initial information needs (Kantz et al., 2025b). We have shown that the extraction of the query can be achieved with a high accuracy, even using small, responsive models (Kantz et al., 2025a). Integrating and evaluating this query system, however, requires further inspections of the ability of users to interact with both the Large Language Model (LLM) and visual query editor. To verify the capabilities and compare it to other systems, we design and conduct a small-scale user study. Within that study, participant had to solve tasks using the combination of the LLM input and graphical query editor. This section of the Deliverable will therefore outline the user study conducted to evaluate our approach, and present the results gathered within this study.

3.1 Related Works

Prior visual querying approaches focus on providing approachable systems to facilitate the use of knowledge graphs for non-experts. GRAPHITE (Chau et al., 2008) and VISAGE (Pienta et al., 2016) both employ visual editing of the prototype graphs to foster the exploration of KGs with a focus on fuzzy matching, query suggestions, and fast retrieval times. RDF Explorer (Vargas et al., 2019) approaches this task similarly by providing a graph editor, enabling the creation of graph examples to retrieve matching instances from the knowledge graph. The system provides users with query expansion options throughout the application, displaying dynamic results as queries are built. They also evaluate their approach using a comparative study, where tasks are given to the users, recording their completion time, progress, and usability using the System Usability Score (SUS) (Brooke et al., 1996). The authors, furthermore, relate their work to previous query builders, notably Smeagol (Clemmer & Davies, 2011). This alternative follows similar exploration and retrieval paradigms but does not offer a comparable number of SPARQL Protocol and RDF Query Language (SPARQL) features. KGVQL (Liu et al., 2022) de nes a novel visual query language to ease the transformation between the visual querying and result set, at the cost of disregarding explorative approaches, favoring the proposed transformation approach to convert between data examples and queries. Rhizomer (García et al., 2022) approaches the exploration of knowledge bases by providing different user interfaces to explore only at the top-level, graph-level, or only at the instance level of a single type. Sparnatural (Francart, 2023) simpli es many of these exploratory approaches into a tree-based approach.

3.2 De nitions

Our KG querying system builds upon the notion of a prototype graph $G_p := (N_p; E_p)$ that serves as the blueprint for all retrieved instances. This graph representation is similar to Basic Graph Pattern (BGP), but it's adhering to the ontology. The ontology constraints the BGP to match to the ontologies classes C and links L. The sub-graph G_p of the ontology consists of the nodes N_p, each one a class C, and edges E_p, each one a link of link types L. The multiset of nodes can contain a class multiple times, and a link can only span the allowed end and start types (within the class hierarchy), i.e.

$$
\begin{aligned}
N_p &:= \quad n_1; : : : ; n_m \quad C \ ; \\
E_p &\quad e_i = (n_t; l_j; n_h) \quad \text{subtypeof}(n_t; \text{fromtype}(l_j)); \\
&\quad \text{subtypeof}(n_h; \text{totype}(l_j)) \quad (N_p \ L \ N_p) :
\end{aligned}
$$

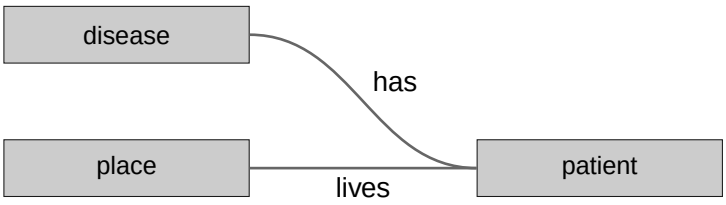

Figure 4: A simple example of a prototype graph $G_p := (N_p; E_p)$ with three nodes and two links.

The function subtypeof(;) takes two nodes and returns a truthy value if the rst one is a subtype of the latter. The fromtype( ) and totype( ) operator take an edge and map them to their node they are from or pointing to, respectively.

An example for the above prototype graph can be found in g. 4, where a person is linked over two edges to a disease and a place. The prototype graph would therefore be

$$G_p = (\{\text{person}, \text{place}, \text{disease}\}, \{(\text{person}, \text{lives}, \text{place}), (\text{person}, \text{has}, \text{disease})\})$$

These prototype graphs G_p represent a query directly, where the query engine matches this prototype graph to an actual instance within the KG.

3.3 User Interface

Our user interface within OnSET provides the users now with an integrated interface to build such a prototype graph. We provide a NLI for users to start the building process by entering a query, which is then translated into a visual query for further re nements in our Visual Query Builder.

3.3.1 NLI

The NLI provides the users a query interface to build visual queries from natural language. We process the query using a three-step-pipeline,

- by rst generating an unconstrained prototype graph from the query using a prompted LLM,
- then retrieving semantically similar classes and links,
- and nally generating the prototype graph again; but with the constraints applied to the LLM retrieved in the previous steps.

This interface is illustrated in g. 5.

We show that this retrieval process is quite robust within Deliverable 5.5 (Schreck, Lengauer, Waldert, Kantz, Boytcheva, et al., 2025), using reference queries sampled from the ontology and comparative outputs. This deliverable also provides further details on the pipeline on the query interface.

3.3.2 Visual Query Builder

Our user interface allows the users to re ne their initial queries in the next step using a node-based editor, where links and nodes can be added to or removed from the prototype graph, all within the constrained link set L, and class set C.

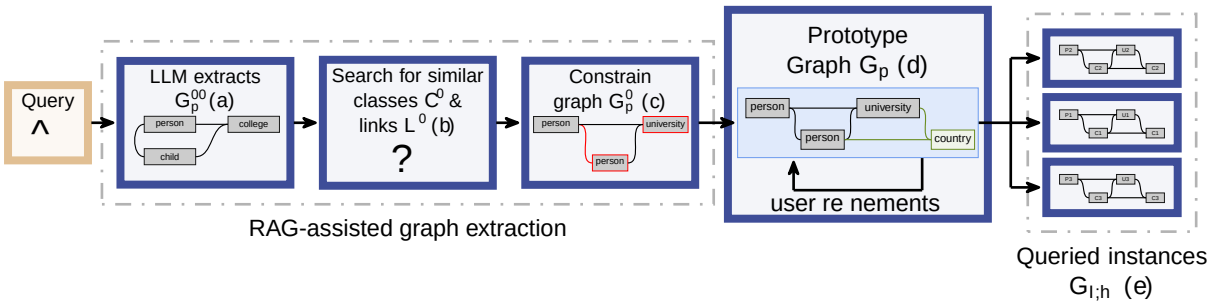

Figure 5: Our query extraction process using LLMs, using the query example “a person and the child of a person have the alma mater of the same university”. We transform the natural language query into a prototype graph using a constrained LLM. The graph is rst approximated using a LLM (a), where the generated classes and links might not match the ontology yet. This initial guess of the LLM is used to search for semantically similar relations (b). With the subset of all possible links and nodes, the graph is extracted again and corrected for possible errors (c), resulting in a graph that adheres to the ontology. The resulting graph can be edited (d), e.g. an additional constraint for a country can be added. The resulting prototype graph can be used to perform queries over a KG to retrieve instances (e).

This re nement can be helpful if our pipeline fails to nd the correct graph or if the users want to re ne their search without an additional text prompt. We support a similar re nement process as the second step in our pipeline, which involves performing semantic search for link retrieval. We use the cosine similarity of sentence embeddings (Reimers & Gurevych, 2019) to compare possible new links that the user can search through when adding these by embedding the query using the same model. The query editor is based on the prior work within the OnSET (Kantz et al., 2025b) and shown in g. 6. The editor supports the creation the prototype graph by allowing the construction by dragging the node-links, with the ability to traverse the class hierarchy. We also enable users to add speci c constraints or query for the label of an instance.

Furthermore, we transparently display how the queries are built using the LLM. This visualization should hold our system accountable for any mistakes and errors that occur during processing and may guide the user towards specifying more precise queries or exploring other querying avenues.

3.4 Study Design

Using this user interface, we test the users on their ability to create queries. We test this by providing our user with multiple tasks, and checking whether they complete them. We additionally ask the users for their rating on a SUS (Brooke et al., 1996).

3.4.1 User Tasks

Our study focuses on two aspects of the user's experience:

1. The user's ability to formulate the correct queries given increasingly complex tasks, with the option to adjust queries using our proposed user interface.
2. The effectiveness of the integrated NLI and visual editor.

Figure 6: OnSET user interface (Kantz et al., 2025b) showing a query. The user can edit the prototype graph G_p on the left, view the connected nodes at the bottom using a circle-packing visualization of the ontology, and inspect the results on the right within the interface.

Table 1: User tasks on the DBpedia KG.

Task	Description	k
Task 0	Find all works where the author is a hockey player.	2
Task 1	Persons that are authors of a work and are gold medalists in a sports event.	3
Task 2	Ships that have their home port in a place in the country United States and persons who have the same place as their death place.	4
Task 3	A work that is the opening theme for a TV show is composed by a person and that person has a child. The person has the alma mater of an University.	5

Table 2: Task results for the OnSET user study on DBpedia with n = 11.

Task	Nodes Correct Rate	Links Correct Rate	Fully Correct Rate	Time (mm:ss)
Task 0	1.00	0.91	0.73	03:46
Task 1	1.00	1.00	0.91	01:12
Task 2	1.00	0.82	0.73	04:49
Task 3	1.00	1.00	1.00	02:05

To validate our claims, we first task the users with completing three increasingly complex queries, as shown in table 1. Each user watches a 2-minute introductory video and has 20 minutes to perform all four tasks. We time the completion of each task and note whether it has the correct number of nodes and links, as well as the correct node types, link types, and constraints. Note that even with fully correct nodes and links, the graph could still not be fully correct due to missing constraints. After they complete all tasks or run out of time, the user is presented with a SUS (Brooke et al., 1996) evaluation questionnaire to estimate their load on our system. We ask the users more general tasks from the DBpedia ontology (Lehmann et al., 2015) to prove the applicability of the system beyond the medical use case.

3.5 Study Results

The user study was conducted at our Institute in Graz by other Computer Scientist. None of the participants had prior expertise in KG querying, nor using the OnSET. We surveyed 11 participants, all with a computer science background, and having completed at least a master's degree.

We report our averaged results in table 2, by task. We observe that all users manage to get all nodes correct. The users struggle with creating the correct links for some instances, however. A fully correct graph is only reached by few users. The average completion rate is still above 70 percent, and rises the more tasks are completed. This could be due to a familiarization effect of the users after spending some time with the system. Nonetheless, all users finish the tasks within the 20-minute limit.

3.6 Discussion

We observed high rates of correctness and fast completion times compared to existing systems like the RDF Explorer, where users took longer to complete their queries while having lower completion rates (Vargas et al., 2019). The results are, however, not directly comparable as they used different query tasks, and different

metrics. This suggests that combining a precise and ef cient LLM system with a visual query editor can reduce the time spent on creating queries while improving the correctness of the results compared to prior works.

4 Talk to Your Graph: Natural Language Querying of Neurodegenerative Diseases

Natural Language Querying (NLQ) and Visual Querying (VQ) serve as accessible methods for working with KGs, both converting user-friendly inputs—whether textual or visual—into the underlying semantic graph structure. TTYG is a technology providing the end user with an interface to communicate with a KG in their natural language. This eliminates the requirement for users to write SPARQL queries or possess expertise in speci c technologies, semantic frameworks, or ontological systems. By enabling users to communicate in English (or other natural language), rather than query languages like SPARQL, this approach reduces technical barriers for non-expert users. TTYG fundamentally changes how people engage with KGs by enabling intuitive access to complex data while remaining independent of the underlying schema and resilient to subsequent modi cations in the KG's organization.

An initial attempt to provide such functionality within the HEREDITARY project was made by delivering the TTYG tool described in Deliverable 5.5 (Schreck, Lengauer, Waldert, Kantz, Boytcheva, et al. (2025)). The interface is built on a speci c KG constructed from open datasets of Ontotext's Linked Life Data Inventory (LLDI)[1] and targeting gene-disease associations. Exploring the knowledge encoded there helps in understanding disease biology (related biological processes, mechanisms, tissue/organ/cell pro ling, interacting partners, etc.), iand denti cation of novel biological targets of new potential drug agents. It can serve to answer questions such as "Which genes are associated with DISEASE-X?", "Which are the biological processes related to DISEASE-X?", "Is there a substance/drug that can treat DISEASE-X?". Now this demonstrator serves as a foundation to validate the user expectations and collect feature requirements for the next iterations of development. Therefore, we focus the current Deliverable around this system and its evaluation by a designed focus group of users. The user feedback is used to collect functional/non-functional requirements, which could be used at later stages of the project to upgrade the existing tools to better t user needs.

4.1 Methodology

Two main methods are used for requirement analysis. The rst method is a literature review, which allows for extensive and deep analysis of existing similar systems and provides good guidance on what are the general necessary requirements. The second method we employ, is a focus group where we give the system to a group of experts for each to test and analyse it, and give feedback based on their experience. We assume that the feedback from experts should give good guidance on what requirements the system should ful ll.

4.2 Literature Review

Recent studies investigate the use of LLMs for natural language querying of RDF KGs. Both and Perevalov (2024) show that LLMs can effectively convert SPARQL queries into verbal descriptions across various languages, making queries more comprehensible for non-technical users. Avila et al. (2024) presents AutoKGQA, a framework leveraging LLMs to convert natural language questions into SPARQL queries. Through querying a KG, the model is able to produce answers based on the results it has extracted from the KG. In the

[1] https://www.ontotext.com/solutions/healthcare-and-life-sciences/linked-life-data-inventory/

biomedical eld, automated KG construction from medical literature combined with LLM-powered interaction methods support clinical decision-making processes and facilitate knowledge discovery (Jadhav et al., 2025).

Chhikara et al. (2025) demonstrates through mem0 that KGs can be used in a Graph-RAG setting as agent memory, allowing LLMs to have a graph that they can query to retrieve interactions and facts, while this functionality is hidden from the user. The use of KGs as a backbone for complex question answering systems has been demonstrated by Markovic et al. (2025). Their system allows LLMs to search through the KGs knowledge and nd provenance for answering complex questions. Recent studies demonstrate the importance of Graph-RAG as superior to standard RAG for retrieval and question answering, as demonstrated by Gutiérrez, Shu, Qi, et al. (2025) and Gutiérrez, Shu, Gu, et al. (2025).

The requirements selected from the literature review are listed along with the user-de ned requirements in Tables 3 and 4.

4.3 Focus Group

To determine what the functional and non-functional requirements are, we created a focus group that consists of data scientists, biomedical domain experts, KGs experts, and AI researchers - a variety of specialists matching the potential users of the tool in the context of the HEREDITARY project. To test the system, we provided them with the standard TTYG UI shown in Figure 7 - a chat interface which allows the user to interact with the system, thus with the KG, in natural language and obtain natural-language responses accompanied by agent-generated explanations (e.g., the rationale behind query construction).

The participants were rst introduced to the system with a short feature overview. Then they were provided with 5 questions (equal for each one of them) with increasing complexity, which they had to execute through the interface. After completing the tasks, they shared their impressions and proposed additional requirements that could enhance their overall user experience with the system.

In total, ten experts took part in the study, most of whom held senior-level positions. Their experience and domain knowledge were essential for gathering meaningful insights into the requirements the system must ful ll.

4.4 Results

The focus of the study is the user experience, which covers mostly functional but also some non-functional requirements. The collected feedback is organized in several dimensions - literature review vs user de ned on one axis. On a different axi,s they are split into must vs nice-to-have ones. While the rst criterion matches the source of the requirement, the second criterion re ects the level of necessity of a given requirement, as reported by the users in the focus group. This separation is important to consider at implementation time, must requirements have higher priority than the nice-to-have ones. In addition, the requirements are classi ed as already implemented, some of which need improvement or new features which are related to not yet implemented but desired functionalities.

4.4.1 Must

The necessary requirements, as determined based on the literature review and the focus group, are listed in Table 3. The last three columns indicate if the requirement comes from the literature review or the focus group, and whether it is already met by the system or will serve for future improvements.

Currently, the persona selection mechanism is not yet met by the system, and that is why it is marked as a new feature. Requirements related to support of different LLMs (APIs), proper Markdown rendering, more simpli ed yet structured information representation, system speed, and, of course, system accuracy are

Figure 7: Figure of the TTYGs UI

already met by the system to a given extent however, they still need improvement in order to provide better user experience and engagement.

4.4.2 Nice-to-Have

The nice-to-have requirements (also referred as may hereafter) are determined based on the literature review and the focus group and are listed in Table 4. The last two columns indicate the source of the requirement and it status regarding implementation - "implemented", "needs improvement", "new feature".

Two requirements of the nice-to-have list are already met (1 and 2); they are related to providing contextual bread-crumbs - the model is prompted after each question to ask relevant questions, and the system is able to verbalize SPARQL queries in natural language. Then three of the requirements are already implemented; however, they need some improvement, and these are related to auto-suggest features, ef cient use of the screen space, and question logging and rating by the users. Four requirements are listed as new features for further development. These are related to the content explainer with graph view, graph zoom functionalities, and taxonomy/ontology browser. These features will again contribute to the user experience, but they will also enable deeper analysis and extended functionality towards downstream tasks.

4.5 Discussion and Future Work

Due to advancements in information extraction, thanks to emerging technologies like LLMs and Graph-RAG, we are closer to allowing domain experts access and query complex knowledge bases and data stores without the need for highly technical skills. TTYG allows for user to query knowledge bases through NLQs.

Table 3: Must requirements.

#	Feature Description	Literature Review	Focus Group	Implemented / New Feature
1	The system must provide accurate and provable answers	X	X	needs improvement
2	The system must be fast to facilitate a good user experience	X	X	needs improvement
3	The system must provide a single conversation sequence with natural language query steps		X	implemented
4	The system must provide complex information in simple ways while avoiding technical complexity		X	implemented
5	The system must provide summarization capability for overall conversations and threads		X	implemented
6	The system must provide session management with the ability to retrieve previous chats, threads, and questions after login		X	implemented
7	The system must provide support for different LLMs models and longer context for instructions and ontology		X	needs improvement
8	The system must provide proper Markdown rendering for improved content presentation		X	needs improvement
9	The system must provide improved user experience with simpli ed yet structured information representation		X	needs improvement
10	The system must provide LLM con guration storage with the ability to switch between different prede ned persona-based con gurations (domain expert vs. non-domain expert)		X	needs improvement
11	The system must provide a persona selection mechanism (supporting different user groups such as standard users and expert users)		X	new feature

The core requirements de ned by the users and based on the literature review are already implemented by the system. However, this study has shown that there is space for improvement of some functionalities, as well as that there are new features that could contribute to a better user experience and engagement. They would also provide facilities for deeper analysis while using the tools and in the subsequent downstream tasks. In summary, improvements are needed in the visualization of the results - especially towards graph views, the ability to use different LLMs interchangeably, and persona-based con gurations. It is interesting to note that no requirements regarding answer provenance were mentioned; our assumption is that the system deals successfully with this important requirement.

This analysis is part of the process of iterative development of visual querying components in HERED-

Table 4: Nice-to-have (may) requirements.

#	Feature Description	Literature Review	Focus Group	Implemented / New feature
1	The system may provide contextual bread-crumbs (schema explainer) showing relation-ship paths between entities		X	implemented
2	The system may be able to verbalize in nat-ural language SPARQL queries which are in-termediate steps of the RAG execution	X		implemented
3	The system may provide auto-suggest fea-tures, including entity name autocomplete and intent suggestions		X	needs improvement
4	The system may make ef cient use of screen space, particularly for graphs and charts with-out wasting whitespace		X	needs improvement
5	The system should provide question logging with ratings (thumbs up/down) for bot im-provement and Q&A dataset collection		X	needs improvement
6	The system should provide multiple result vi-sualization formats, including text output, ta-ble representation, and graph visualization		X	new feature
7	The system may provide content explainer in a separate tab with natural language and graph view		X	new feature
8	The system may provide graph zoom func-tionality supporting class/instance level nav-igation		X	new feature
9	The system may provide a taxonomy/ontol-ogy browser for exploring entity relationships		X	new feature

ITARY project. It will serve as a reference point for further developments and improvements in the next period. Depending on the project development and partners needs, the improvements and new features will be further discussed, prioritized and where appropriate - brought to implementation. They may also spike the development of new tools which would ful ll better some of the user needs.

5 Concept for Visual Analysis of Cell-Cell Signaling

In medical applications, the increasing availability of single-cell RNA sequencing data from pathological and healthy samples has led to the development of methods for differential analyses, i.e. extensions of classic bioinformatics analyses which identify signi cant biological alterations in the pathological condition with re-spect to an healthy reference. These methods now also comprehend the elds of Cell-Cell Communication (CCC) inference and Transcription Factors (TF) activity analysis (Cesaro et al. (2025)), which provide infor-mation on disrupted molecular signaling processes in the form of high-dimensional results that need to be

interpreted by biomedical experts of the disease under study.

The results of differential CCC and TF analyses can be represented as multi-layer networks. Each multi-layer network is de ned for the combinations of sender-receiver cell types interacting in the data set (see Figure 8a). These networks consist of three layers: i) sender intra-cellular signaling involving TFs and ligand-encoding target genes, Fig. 8a upper layer; ii) inter-cellular differential communication involving sender-expressed ligands and receiver-expressed receptors, Fig. 8a mid layer; iii) receiver intra-cellular communication involving receptors and downstream TFs, Fig. 8a lower layer. Each layer represents molecules as nodes, with edges denoting molecular or regulatory interactions among them. Depending on the layer, edge weights represent different biological properties: the activating or inhibitory activity of TFs on ligand-encoding target genes, the magnitude and signi cance of differential CCC events, or the existence of a signal transduction cascade linking receptors to downstream TFs.

Due to the complexity and size of CCC and TF results, interactive visualizations provide a valuable strategy for a time-ef cient identi cation of key molecules and signaling pathways that drive disease mechanisms.

5.1 Multi-layer Network Visualization

In the context of multi-layer network visualization, McGee et al. (2019) provide a categorization of the eld based on the task to be performed, while Filipov et al. (2023) formulate a roadmap based on network characteristics such as categorical attributes and spatiotemporal information. In the differential cell signaling case study a central analytical goal is the identi cation of disrupted signaling loops, which involve TFs and ligand-receptors interactions affecting each other's activity, hence creating a self-sustaining loop of over-/under-activation in the cell signaling paths. It can be thus considered as an instance of task category A of McGee et al. (2019), which focuses on cross-layer entity connectivity tasks. Moreover, the network includes both categorical and numerical attributes on nodes and edges, it can thus be characterized as a group multi-variate network according to Filipov et al. (2023) roadmap to network visualization, i.e. a network that includes different facets, where facets are data attributes belonging to nodes, edges, or the entire graph, that need to be visualized along with the graph's topology. For a comprehensive interpretation of the pathological alterations, in addition to signaling loops, it is also essential to identify molecules, i.e. nodes, having a prominent impact on cell signaling mechanisms, since these molecules could be candidate drivers of the pathology. Metrics such as betweennes and centrality, adapted to weighted, cross-layer paths, can provide quantitative estimates of node importance (Ehlers et al. (2025)) and guide user's interpretation of results.

In order to facilitate network interpretation, visualizations will be enhanced with speci c user interaction techniques. Graph lenses provide an effective interactivity strategy to let the user focus on speci c sub-graphs (Tominski et al. (2006)) of interest within cluttered visualizations, providing an integrated representation of neighbors and overall network structure. Interactions like hovering and clicking will be used to highlight key features of the elements selected by the user.

5.2 Dashboard Design

The dashboard will address domain speci c requirements such as i) user-friendly query de nition, to focus on speci c cell types or molecules, ii) facilitated identi cation of molecules leading the pathological signaling mechanisms, through network science metrics, iii) possibility to upload a custom dataset, to visualize and interpret new case studies, iv) visualization layouts that highlight signaling loops, v) summary metrics on sender-receiver networks. The proposed differential cell signaling dashboard (Fig. 8b) has a structure that aims to maximize exibility, through user-de ned queries, several visualization layouts and user-de ned uploaded case studies, and to improve biomedical interpretability, in particular, through the integration of an LLM-based chat assistant that facilitates the retrieval of literature about speci c molecules and diseases. A

Figure 8: Graphical illustration of (a) cell-cell signaling conceptualization and modeling through a multi-layer network, (b) dashboard design, and (c) schematic example of biological hypothesis drawn from multi-layer network interpretation. Created with BioRender.com

Figure 9: Screenshot of a rst draft of the dashboard, showing an initial implementation of the concentric circular layout.

left-hand sidebar will allow users to: (i) select between a pre-computed case study, choosing between ALS and FMD analyses, or upload custom CCC and TF results (CSV format), and (ii) lter the network using numeric thresholds on node- and edge-level attributes. The main section will provide a comprehensive overview along with several cell types - speci c visualizations. Users will rst be able to inspect the overall multi-layer network, with graph lenses interactivity, and subsequently focus on signaling interactions involving speci c cell types of interest. Alternative layout algorithms for node placement will be included to provide different representations of the signaling alteration: tree, concentric circular, linear, 2.5D and hive options will be provided. In addition, techniques such as edge bundling will be employed to reduce visual clutter and emphasize common interactions. As a conceptual example, in 8b bottom-right, a visualization having nodes placed on concentric circles de ned by categorical variables is represented. To provide also simpli ed and focused visualizations, a panel will be dedicated to plots of subsets of nodes, derived from CCC or TF data. A right-hand sidebar will integrate an LLM-based scienti c assistant designed to retrieve and summarize relevant literature associated with user-selected molecules, thereby supporting biological interpretation and hypothesis generation.

Together, these components will form an integrated framework that substantially reduces the time and effort required for domain experts to interpret single-cell computational results, enabling faster identi cation of dysregulated cell–cell signaling events and more ef cient generation of biological hypotheses. A rst draft of the dashboard is shown in 9.

5.3 Next Steps

This visualization concept is based on a survey of visualization techniques for multilayer networks, and collaboration with researcher partners. We are in the process of implementing this visualization and interaction design into a prototype application, which in turn we can evaluate with stakeholders for further development.

6 Initial Design of a User Study on Glyph Design

As part of Deliverable 5.3 (Schreck, Lengauer, Waldert, Kantz, Grabner, et al., 2025) a novel glyph design was developed for a hybrid data modality combining both proportional and spatial properties. This kind of data often results from spatial transcriptomics, a technique to capture gene expressions while still preserving their spatial context (Miller et al., 2022). The incentivization of the design stems from the Bio+MedVis Challenge[2] @ IEEE VIS 2024, where tissue sample of a coronal cut of a mouse brain was provided together with “spot-wise” gene expressions. Standard visualization techniques such as pie charts can be displayed directly on top of the tissue sample as an overlay at the respective locations (Figure 10c). However, this kind of solution renders it very hard to make out similarity clusters, as pie charts are ill-suited for between-subject comparison (Tufte & Graves-Morris, 1983). As a solution, we developed and published the Dense Representation of Proportion Landscapes Enhancing Trends (Droplet) glyph (Lengauer et al., 2024), which splits up the individual components of the pie chart and renders them as separate, yet tightly packed shapes with sizes relative to their respective contributions. A circular base shape is equipped with a protruding tail (i.e., forming a droplet shape), pointing in the direction of the closest belonging cluster. Even though it appears that this novel glyph type allows us to perceive similarity patterns more easily (Figure 10d), their actual effectiveness and ef ciency for solving speci c visual analytics tasks remains untested. That remains to be established through a dedicated perception-based user study.

[2] http://biovis.net/2024/biovisChallenges_vis/

(a) Glyph Concept

(b) Mouse Brain Example

(c) Closeup with Pie Charts

(d) Closeup with Droplets

Figure 10: Droplets glyph design (a) showing spatial transcriptomics data of a coronal cut of a mouse brain, superimposed to a Hematoxylin and Eosin stain (b). A side-by-side comparison of a baseline pie chart visualization (c) with our solution (d) underlines the bene t of the latter.

6.1 Related Work

The Droplet glyph is novel in its design, but there are related concepts for visualizing proportional relations, such as Stixels (Mörth et al., 2020), or Vizent Glyphs (Holliman et al., 2024). The idea of indicating trends is