

(En-)Coding The Knowledge Graph Stack: RDF, RDFS, SPARQL, SHACL, (YARR)RML^{*}

Achim Reiz^{1,*}, Gregor Seifer¹

¹neonto GmbH, Luxemburger Straße 264, Cologne, Germany

Abstract

Knowledge Graphs help break down data silos by connecting heterogeneous sources into a shared, machine interpretable data layer. Too often, the stack driving this technology remains a mystery. This hands-on tutorial introduces the underlying technologies for building RDF-based knowledge graphs. In two sessions, participants will learn to understand and create graphs, apply ontologies, build constraints and ingest data.

The first half of the day, we will start with an introduction into triples, how they can be built with a simple text editor and how they already form a graph. Some of the terms carry fixed semantics for building ontologies (RDFS, OWL) or are from already established vocabularies, and we will explore how to use them in a practical setting. Later in that session, we explore SPARQL for extracting our newly created data.

In the second half of the day, we jump into ingestion and constraints. First, we create a data ingestion pipeline using YARRRML, a simpler subset of R2RML and RML, the standardized languages for creating declarative pipelines. Later, we will apply data quality checks with SHACL, ensuring that our graph data is structurally consistent, thereby supporting a high degree of data quality.

Keywords

RDF, RDFS, SHACL, SPARQL, YARRRML

1. Why This Tutorial?

1.1. The Goal

Too often, semantic web technologies are seen as something that is hard to manage and learn. In this session, we would like to show that we do not need expensive, specialized software to deal with graphs. It is enough to have the text editor and some understanding of the tool stack.

1.2. Why Now?

This course teaches the fundamentals of knowledge graphs and the semantic web. While we believe this is an almost timeless skill in general, its importance increases in the realm of AI. Generative AI systems cannot be built without a grounding in structured, well-defined data, and both research and industry point to knowledge graphs as the prime structuring tool [1]. This course teaches the fundamentals on how to build these knowledge graphs.

1.3. Why Semantics?

SEMANTiCS audiences care about building systems that integrate heterogeneous data, preserve meaning across contexts, and remain interoperable over time. The RDF stack sits at the center of that mission: it provides standard, vendor-independent building blocks for modeling (RDF and RDFS), querying (SPARQL), validating quality (SHACL), and ingesting data at scale (YARRRML/R2RML/RML). By combining these pieces in a hands-on format, the course helps participants move from high-level

Posters, Demos, Blue Sky, and Tutorials at SEMANTiCS 2026, Sep 2026, Ghent, Belgium

^{*}Posters, Demos, Blue Sky, and Tutorials at SEMANTiCS 2026, Sep 2026, Ghent, Belgium

^{*}Corresponding author.

✉ achim.reiz@neonto.de (A. Reiz); gregor.seifer@neonto.de (G. Seifer)

ORCID 0000-0003-1446-9670 (A. Reiz)



© 2026 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

knowledge graph concepts to implementable, maintainable semantic solutions, which directly supports the conference's goal of turning semantics into working systems.

1.4. Why These Speakers?

Achim Reiz combines deep research experience with a clear, practice-oriented perspective on what makes knowledge graph projects succeed. He obtained his Dr.-Ing. at the University of Rostock, working on quantitative methods to assess knowledge graphs using quality metrics. That gave him a rigorous understanding of how to measure and reason about graph quality in a systematic way.

Through this work, he identified a recurring root cause behind many quality problems: Teams often start with the wrong tooling and lack a solid understanding of vocabularies and their downstream effects on interoperability and data clarity. This insight led him to found [neonto](https://neonto.de/)¹, a company focused on teaching the fundamentals and specifics of RDF-based knowledge graphs and providing tooling that supports organizations in bringing interoperable, standardized data structures to life.

Gregor Seifer comes from a practical, cross-domain background. He obtained his M.Sc. in Artificial Intelligence at Ajou University in South Korea and has hands-on experience from multiple data alignment projects in data warehouse and BI initiatives. Since joining [neonto](https://neonto.de/) this year, he has expanded this experience into the knowledge graph space and now works as a consultant, supporting organizations in designing, aligning, and operationalizing knowledge graphs.

2. About the Tutorial

2.1. Structure

We plan to structure the tutorial as a full-day workshop with four sub-sessions on RDF/RDFS, SPARQL, YARRRML, and SHACL. The structure can be condensed to a half-day workshop; then we will skip the session on declarative mappings and condense the rest of the contents.

Table 1
Timetable

Time	Topic	Contents
11:00–11:45	RDF Fundamentals	Triples, graph model, RDF syntax, URIs, literals, RDF Schema, reusing existing vocabularies
11:45–12:30	SPARQL	Querying RDF, graph patterns, filters, aggregation, hands-on querying exercises
12:30–13:30	Lunch break	
13:30–14:15	(YARR)RML	Declarative data mappings, ingestion pipelines, hands-on mapping exercises
14:15–15:00	SHACL	Data validation, node and property shapes, integrating validation into RDF workflows

The tutorial is highly interactive: We will work in a VS Code web environment that comes with the required RDF tooling (SHACL-engine, language server, ingestion engine) and together gain hands-on experience in these technologies.

2.2. Target Audience

This tutorial is designed for knowledge graph beginners and early practitioners who want a clear, practical introduction to the RDF stack without assuming prior semantic web expertise. It is particularly relevant for data engineers, data architects, software developers, and consultants who work with

¹<https://neonto.de/>

heterogeneous data sources and want to understand how RDF-based knowledge graphs can be built, queried, validated, and maintained in practice.

However, this tutorial also addresses AI and data professionals who encounter knowledge graphs as part of integration, analytics, or AI workflows and want to understand the underlying standards rather than treating them as a black box and part of their UI. By combining conceptual foundations with hands-on exercises, the course supports participants who want to move from high-level knowledge graph concepts to confidently creating, ingesting, querying, and validating RDF data themselves.

The tutorial is structured consecutively, starting with the core concepts of RDF and gradually extending them with querying, ingestion, and validation. Each session introduces the concepts needed for the next one, creating a coherent learning path rather than isolated topics. At the same time, the modules are self-contained enough that participants with prior experience can comfortably join at a later point, for example directly at SPARQL, mappings, or SHACL, without having followed every earlier exercise.

2.3. Technical Requirements

Participants will require a notebook with stable internet access to follow along. Ideally, the room provides tables and power sockets.

2.4. Related Tutorials

We held parts of the tutorial at Connected Data London 2024² and 2025³.

2.5. Notes On Proprietary Software Use

With neonto, we develop RDF technologies for use in a programmer-like environment. While we are going to utilize the browser-based variant of our editor⁴, which is part of our RDF academy, this software is mainly developed to be a better interface (IDE) for existing RDF standards. Thus, the gained knowledge is applicable to a variety of closed-source and open-source applications.

Declaration on Generative AI

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

References

- [1] J. Sequeda, D. Allemang, and B. Jacob, "Knowledge Graphs as a source of trust for LLM-powered enterprise question answering," *Journal of Web Semantics*, vol. 85, 2025, Art. no. 100858. <https://doi.org/10.1016/j.websem.2024.100858>

²<https://2024.connected-data.london/talks/building-your-knowledge-graph-with-rdf-an-introduction/>

³<https://2025.connected-data.london/talks/shacl-building-flexible-schemas-for-knowledge-graphs/>

⁴<https://neonto.de/solutions/editor/>