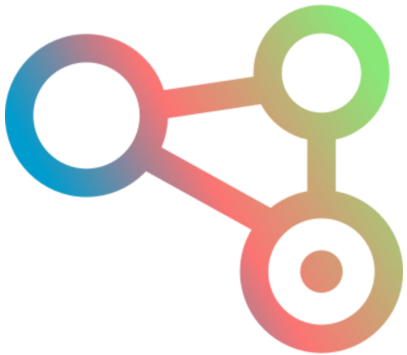
P16 Days



Corese

Pierre Maillot, Research engineer, *INRIA*
https://w3id.org/people/pierremaillot

01

Knowledge graphs: Some facts

Some history

Some history

> Knowledge graphs
  - Early steps in 1975
> Several approaches:
  - Property graphs
  - Semantic web
  - Hybrids
> “Semantic web” coined in 2001 by Tim Berners-Lee

Some semantic web history

- RDF: 22 février 1999
- RDFS: 10 février 2004
- OWL: 10 février 2004
- SPARQL 1.0: 15 janvier 2008
- OWL 2: 12 décembre 2012
- SPARQL 1.1: 21 mars 2013
- SHACL: 20 juillet 2017
- RDF Canonicalization: 21 Mai 2024

What are semantic web graphs ?

Technologies

- Representation of information as triples
  - With RDF, RDFS and OWL
- Query language
  - With SPARQL
- Validation mechanism
  - With SHACL
- Integration in the web
  - Everything in based on HTTP. No new protocol to learn
- Emergent uses
  - RAG, IoT,

Advantages

- Versatility
  - Anything can be described
- Ease of integration
  - RDF + RDF = Still RDF
- Standard-first
  - W3C, an international non-profit consortium, sets the standards
- Reasoning
  - Say less, infer later
- Built-in provenance

Semantic Web

Similarities with SQL databases

> Joint-based data management systems

> Query language

> Validation

Differences with SQL databases

> Age (1980 vs 2008)

> Schema flexibility

> No downtime during schema update

> Better at handling heterogeneous data

Semantic Web - In Use

Multiple offers

> Open-source

- Apache Jena, Eclipse RDF4J, Openlink Virtuoso

> Paid

- Openlink Virtuoso, Ontotext GraphDB

Multiple uses

> Data integration

- IBM, Siemens

> Search, recommandation

- Google, Amazon, Alibaba, Walmart

> Graph RAG

- Microsoft, NASA

02

Knowledge graphs: Corese

What is Corese ?

Semantic Web Factory

Knowledge graph engine started by Olivier Corby in 1999

- Support of all standards
- Exploration of new or underrepresented functionalities
- Tools to facilitate the usage of knowledge graphs

Corese, some history

Support of the standard

- RDF: 22 février 1999✅
- RDFS: 10 février 2004✅
- OWL: 10 février 2004
- SPARQL 1.0: 15 janvier 2008
- OWL 2: 12 décembre 2012
- SPARQL 1.1: 21 mars 2013 ✅
- SHACL: 20 juillet 2017✅
- RDF Canonicalization: 21 Mai 2024✅

But, also

> LDScript
  - Adds programming to SPARQL queries

    Allows HTTP interaction, non-RDF format interaction, loops, etc.

> Custom reasoning engine
  - Adds custom reasoning rules using SPARQL queries

> Federated SPARQL
  - Send part of SPARQL queries to remote endpoints

Corese, technology

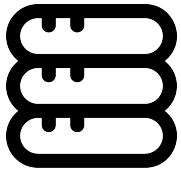
In-memory storage coded in Java

Connectors to other industry solutions for permanence

Corese core

Java

Library



Corese server

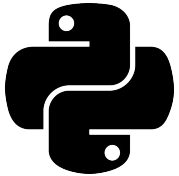
Java

Server

Corese GUI

Java

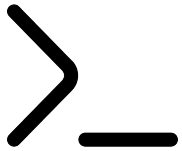
Standalone client



Corese Python

Python

Python binding



Corese Command

Java

CLI

Corese RDF4J/Jena

Java

Connector to persistent storage

Corese - In Use

Research

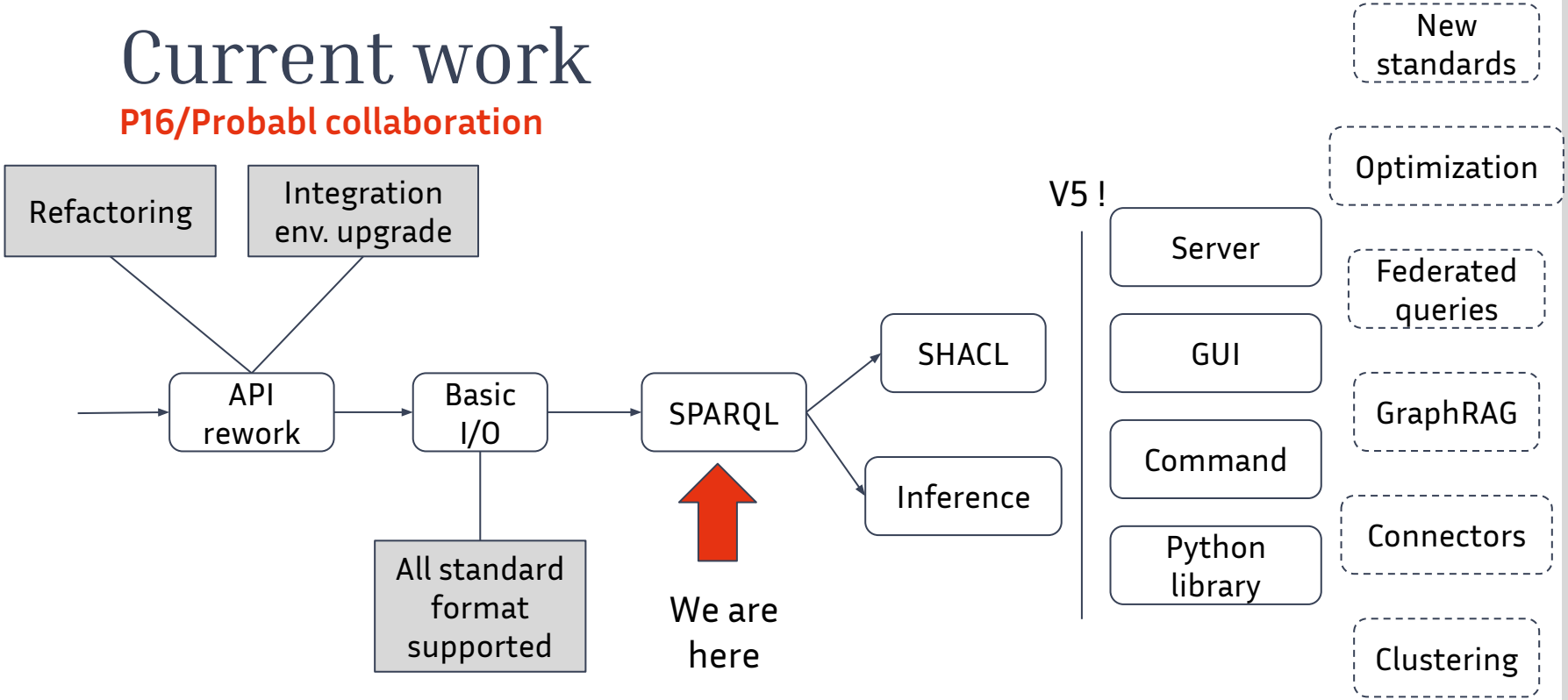
- Extension of SHACL with probabilities Rémi FELIN PhD thesis in November last year
- 2 on-going PhD thesis at WIMMICS
    - Nicolas Robert: Knowledge graph embedding models: symbolic knowledge injection and discovery
    - Célian Ringwald: Learning RDF pattern extractors for a language from dual bases Wikipedia/LOD
- On-going work on UI for the exploration of knowledge graphs by Aline MENIN

Industry

- Connector to Elasticsearch server for Probabl

Current work

P16/Probabl collaboration

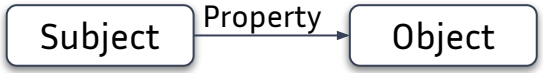


Thank you !

https://github.com/corese-stack

Inner details

Indexing



Quad (Subject, property, object, graph)

Index PSOG
POSG
PGSO

Inner details

Reasoning

Reasoning engine

Based on the SPARQL engine

If SPARQL query have results Then Construct triples