

What are Knowledge Graphs? Why are they Key to Successful Data Governance?

Irene Polikoff, TopQuadrant
Jan Voskuil, Taxonic

February 21, 2019

What are Knowledge Graphs?

Why are they Key to Successful Data Governance?

- INTRODUCTION
 - The history of knowledge graphs in 90 seconds
 - What is data governance
 - What is a knowledge graph, how does it work
 - Why data governance needs knowledge graphs
 - TopBraid EDG as a knowledge graph for data governance

- DEMO

- WRAP-UP

- Questions and Answers



Irene Polikoff



Jan Voskuil

WHO WE ARE



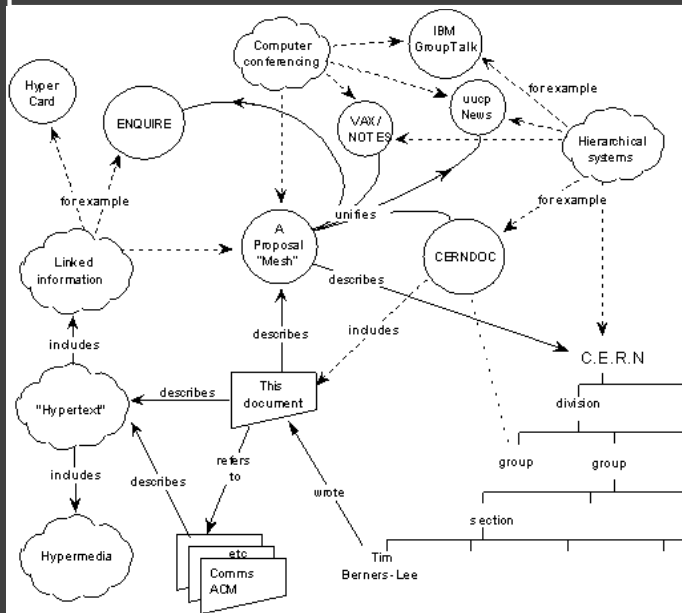
- **Founded:** in 2001 with a strong commitment to standards-based semantics
- **Mission:** Make enterprise data meaningful
- **Focus:** Provide comprehensive data governance solutions



- **Founded:** in 2012 with a strong commitment to model driven architectures
- **Mission:** Use semantics to transform the business
- **Focus:** Implement TopQuadrant's TopBraid suite for local customers

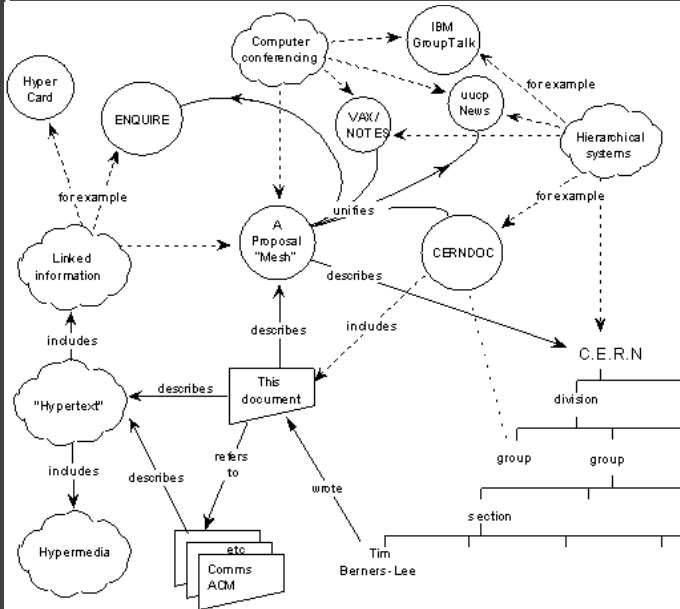
KNOWLEDGE GRAPHS

Tim Berners Lee




KNOWLEDGE GRAPH

Tim Berners Lee



2007 Google








Ada Lovelace

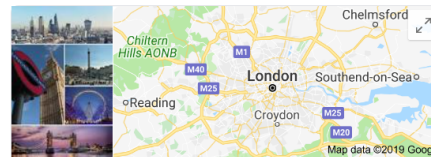
Mathematician

Augusta Ada King, Countess of Lovelace was an English mathematician and writer, chiefly known for her work on Charles Babbage's proposed mechanical general-purpose computer, the Analytical Engine. [Wikipedia](#)

Born: December 10, 1815, [London, United Kingdom](#)
Died: November 27, 1852, Marylebone, United Kingdom
Spouse: William King-Noel, 1st Earl of Lovelace (m. 1835–1852)
Children: [Anne Blunt](#), [15th Baroness Wentworth](#), [MORE](#)
Parents: Lord Byron, Lady Byron
Known for: Mathematics, Computing

People also search for View 15+ more

 Charles Babbage	 Lord Byron Father	 Lady Byron Mother	 Grace Hopper	 Alan Turing
--	---	---	---	--



London

Capital of England

London, the capital of England and the United Kingdom, is a 21st-century city with history stretching back to Roman times. At its centre stand the imposing Houses of Parliament, the iconic 'Big Ben' clock tower and Westminster Abbey, site of British monarch coronations. Across the Thames River, the London Eye observation wheel provides panoramic views of the South Bank cultural complex, and the entire city.

Area: 1,572 km²

Elevation: 11 m

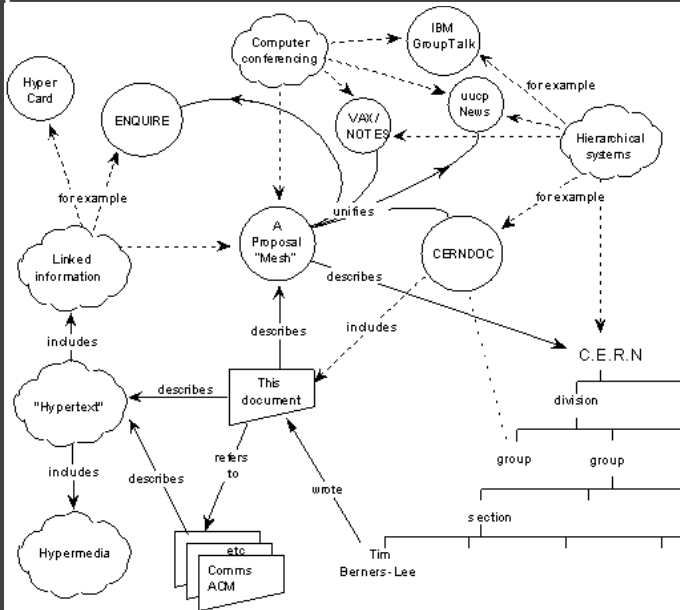
Weather: 8°C, Wind SW at 8 km/h, 87% Humidity

Local time: Wednesday 06:59

Population: 8.136 million (2011) [United Nations](#)



Tim Berners Lee



2007 Google

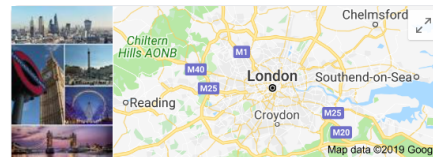
Ada Lovelace
Mathematician

Augusta Ada King, Countess of Lovelace was an English mathematician and writer, chiefly known for her work on Charles Babbage's proposed mechanical general-purpose computer, the Analytical Engine. Wikipedia

Born: December 10, 1815, [London, United Kingdom](#)
Died: November 27, 1852, Marylebone, United Kingdom
Spouse: [William King-Noel, 1st Earl of Lovelace](#) (m. 1835–1852)
Children: [Anne Blunt](#), [15th Baroness Wentworth](#), [MORE](#)
Parents: [Lord Byron](#), [Lady Byron](#)
Known for: Mathematics, Computing

People also search for View 15+ more

[Charles Babbage](#) [Lord Byron](#) [Lady Byron](#) [Grace Hopper](#) [Alan Turing](#)



London

Capital of England

London, the capital of England and the United Kingdom, is a 21st-century city with history stretching back to Roman times. At its centre stand the imposing Houses of Parliament, the iconic 'Big Ben' clock tower and Westminster Abbey, site of British monarch coronations. Across the Thames River, the London Eye observation wheel provides panoramic views of the South Bank cultural complex, and the entire city.

Area: 1,572 km²

Elevation: 11 m

Weather: 8°C, Wind SW at 8 km/h, 87% Humidity

Local time: Wednesday 06:59

Population: 8.136 million (2011) United Nations

2019

Knowledge Graphs in the news

- > Google on AI and knowledge graphs in zdnet
- > Data governance 2.0 on Dataversity
- > Conferences & workshops
- > Technology trends 2019
- > Forbes article by Kurt Cagle



Data governance



- > Data represent value
- > Usage of data is subject to regulation
- > Bookkeeping of data as assets is necessary

“... organizing **people**, **processes**, and information **technology** required to create a consistent and proper handling of **an organization's data** across the business **enterprise**.” ... Wikipedia

How is it done?

- > Ad hoc
- > Documents, spreadsheets
- > Examples:
 - > GDPR
 - > Data quality
 - > Data lineage



▲ AEX +0,80% 623,24
▼ MIDKAP -0,03% 725,56
▲ EUROSTOXX 50 +0,37% 3.171,12
▲ NAS 100 +0,09% 2.708,53
▲ C/S NEW YORK +0,1045% 1.1459
▼ RENTE NL 10JR -0,005 0,250%

EUROPEAN NEWSPAPER OF THE YEAR

fd.

het financieele dagblad

Maandag 4 februari 2019 • Nummer 29 • 32 pagina's
 'Al het geene tot het financieele eenige betrekking heeft' - sinds 1796

Tegenslag
 'Al vanaf de snelweg zag ik de rookwolken opstijgen'

PAGINA 18



ADVERTENTIE

CPI
 Connecting People with Impact
 The future in risk, finance & governance
 Meet James, Partner
 tel. 32 78 97 84
 www.meetcpi.com

Overzicht

VERZEKERAARS

DNB maant Vivat om lakse screening op terrorisme

Voormalig Reaal heeft controles niet op orde

Saskia Jonker en Johan Leupen
 Amsterdam

Vivat heeft een tik op de vingers gekregen van toezichthouder De Nederlandsche Bank (DNB), omdat de verzekeraar zijn interne controlesystemen om terrorismefinanciering te voorkomen al jaren niet op orde heeft. Dat blijkt uit een brief van DNB die het FD heeft ingezien. Vivat, het bedrijf achter de merken Zwitsersleven, Reaal en Actiam, moet vrezen voor een sanctie als het niet snel verbeteringen doorvoert. Ondanks herhaalde waarschuwingen schiet de verzekeraar op sommige vlakken 'ernstig tekort' bij het mitigeren van klanten of andere relaties op sanctielijsten staan, zo kreeg de directie van Vivat schriftelijk te verstaan.

Onduidelijk is zelfs of alle relaties wel zijn gescand, schrijft DNB. Dat had wel moeten volgens de Sanctielijst. Dit was...

1 min
 De maximale boete is in 2017 verhoogd van €10.000 naar €1 mln.

volgens informatie van het FD er toe geleid dat Vivat in allerlei adviseur EPMG heeft ingehuurd om op orde op zaken te stellen. De verzekeraar heeft er alle belang bij zijn systemen snel te verbeteren, omdat het bedrijf sinds half januari te koop staat. Waar de Chinese eigenaar Anbang gaat voor de hoogste prijs voor Vivat, kunnen potentiële kopers de problemen aangrijpen om een korting te bedingen.

Daar komt bij dat de toezichthouder volgens kenners in de markt bij dit dossier zijn tanden wil laten zien. In de ING-zaak werd DNB ingehaald door het Openbaar Ministerie, dat de boete van €775 mln oplegde.

Ook in de internationale arena moet helder zijn dat DNB er bovenop zit, zeker omdat de Amerikaanse justitie er niet voor terugdeinst Europese financiële instellingen miljarden boetes te opleggen.

In dezelfde vijver

Bijna de helft van de vacatures blijft onvervuld. Werkgevers moeten alles uit de kast halen

PAGINA 10



The problem with data governance

- > Galaxies of data
- > Diversity of perspectives
 - > Business
 - > Technical
 - > Regulatory
- > Diversity of representation

Solution

- > Start with business glossaries
- > Link to abstract information models
- > Link to physical data stores



What is a Knowledge Graph?

- > A Knowledge Graph represents a knowledge domain
- > It represents knowledge as a graph
 - > a network of nodes and links
 - > not tables of rows and columns
- > It represents facts (data) and models (metadata) in the same way
- > It is based on open standards, from top to bottom

```
mi6:JamesBond mi6:worksFor mi6:MI6 .  
mi6:JamesBond rdf:type schema:Person .  
mi6:MI6 rdf:type schema:Organisation .
```

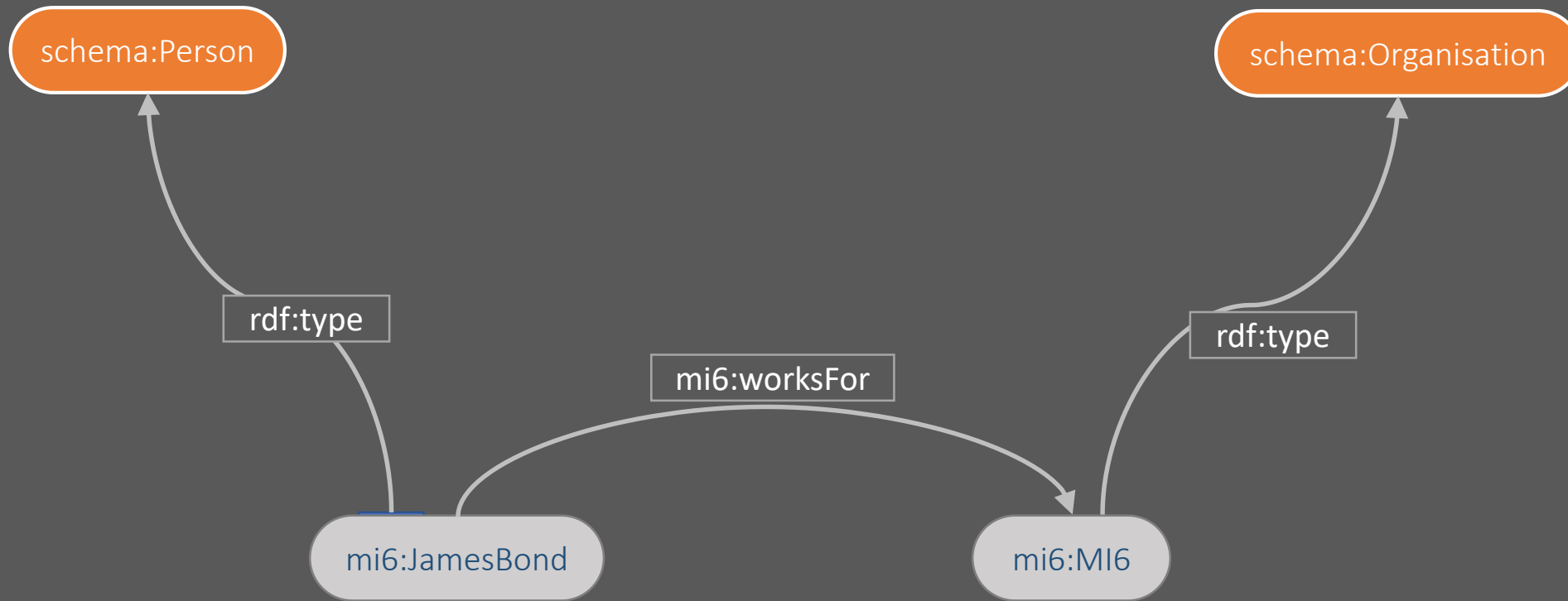
SUBJECT

PREDICATE

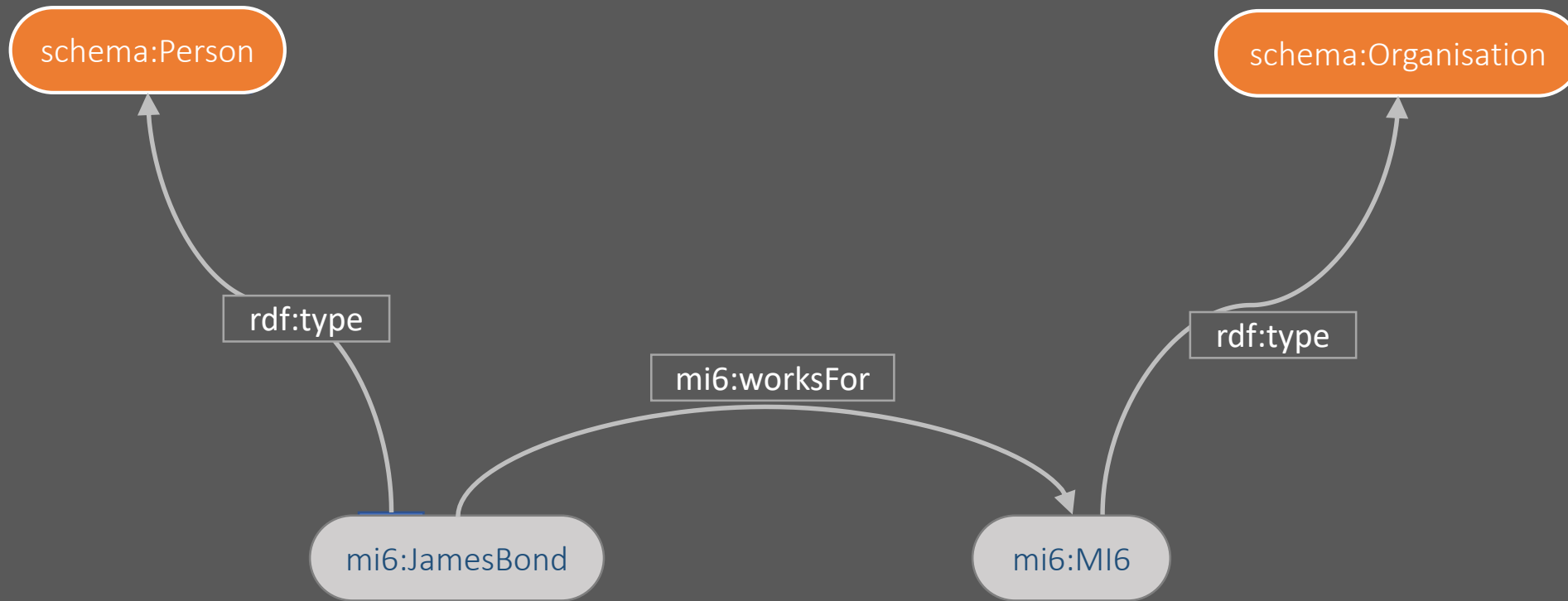
OBJECT

```
mi6:JamesBond mi6:worksFor mi6:MI6 .  
mi6:JamesBond rdf:type schema:Person .  
mi6:MI6 rdf:type schema:Organisation .
```

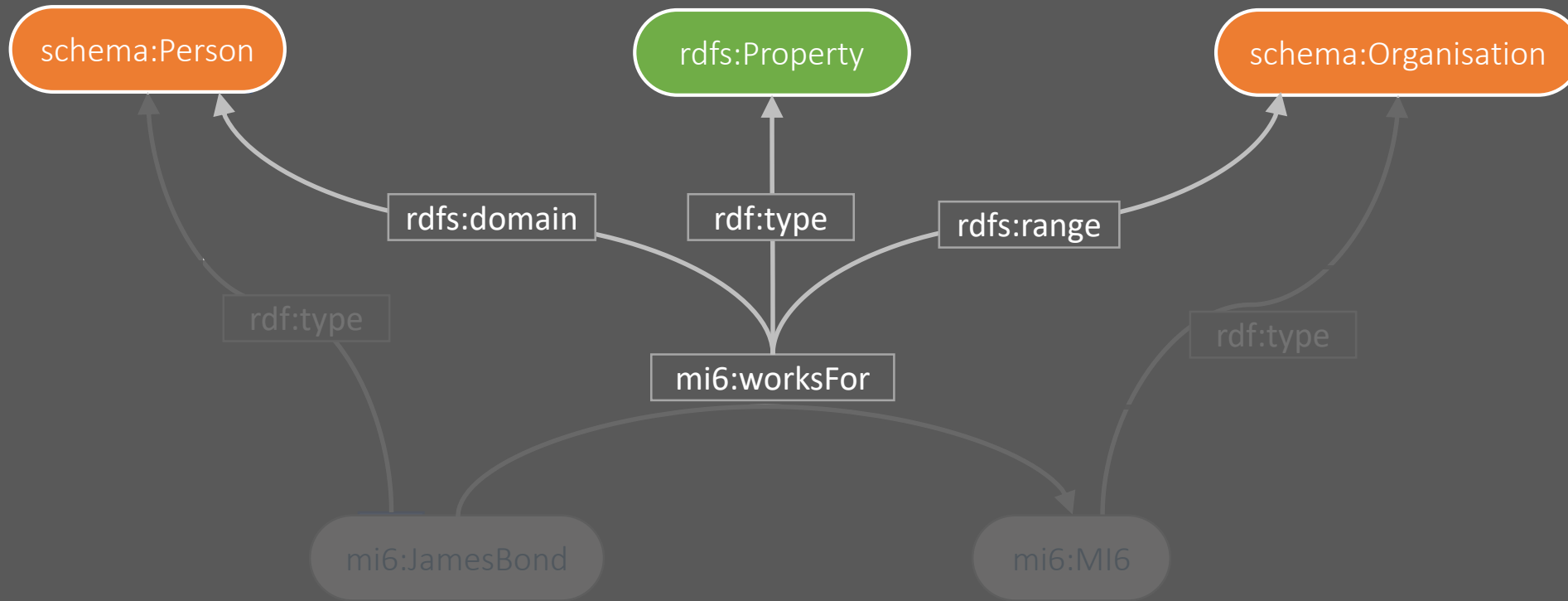
```
mi6:JamesBond mi6:worksFor mi6:MI6 .  
mi6:JamesBond rdf:type schema:Person .  
mi6:MI6 rdf:type schema:Organisation .
```



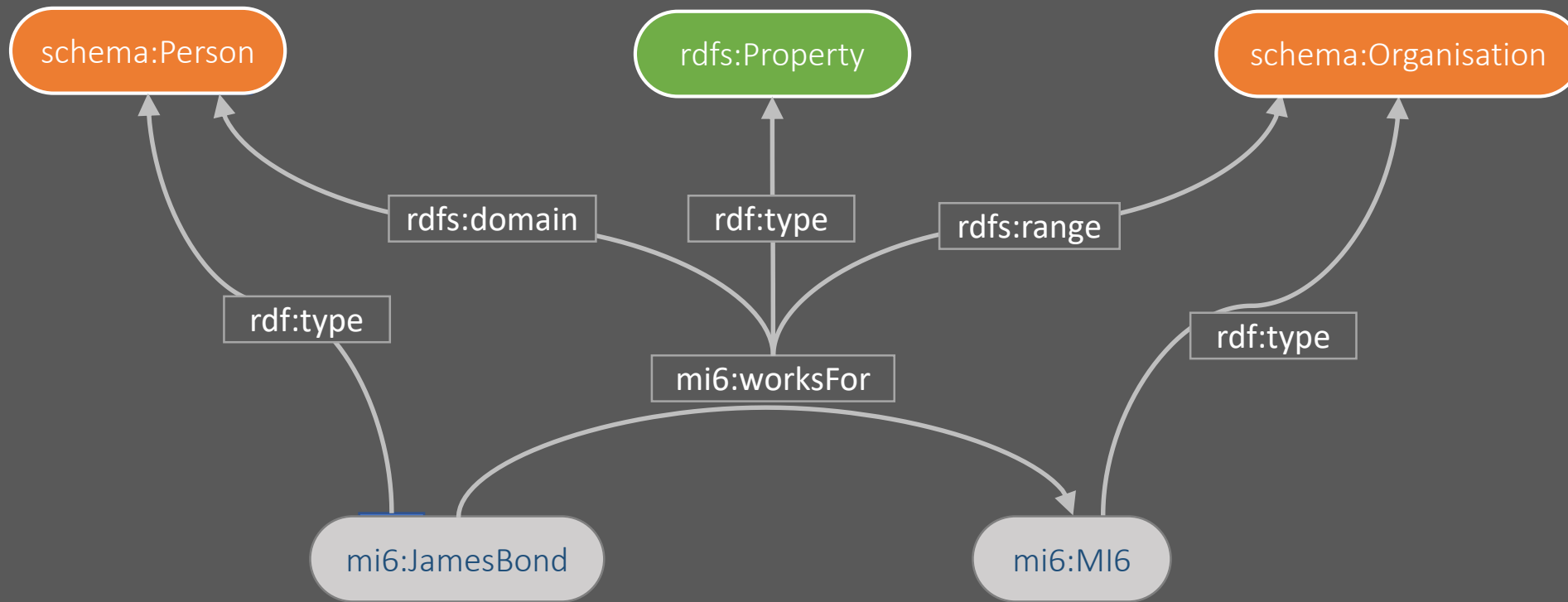
```
mi6:JamesBond mi6:worksFor mi6:MI6 .  
mi6:JamesBond rdf:type schema:Person .  
mi6:MI6 rdf:type schema:Organisation .
```

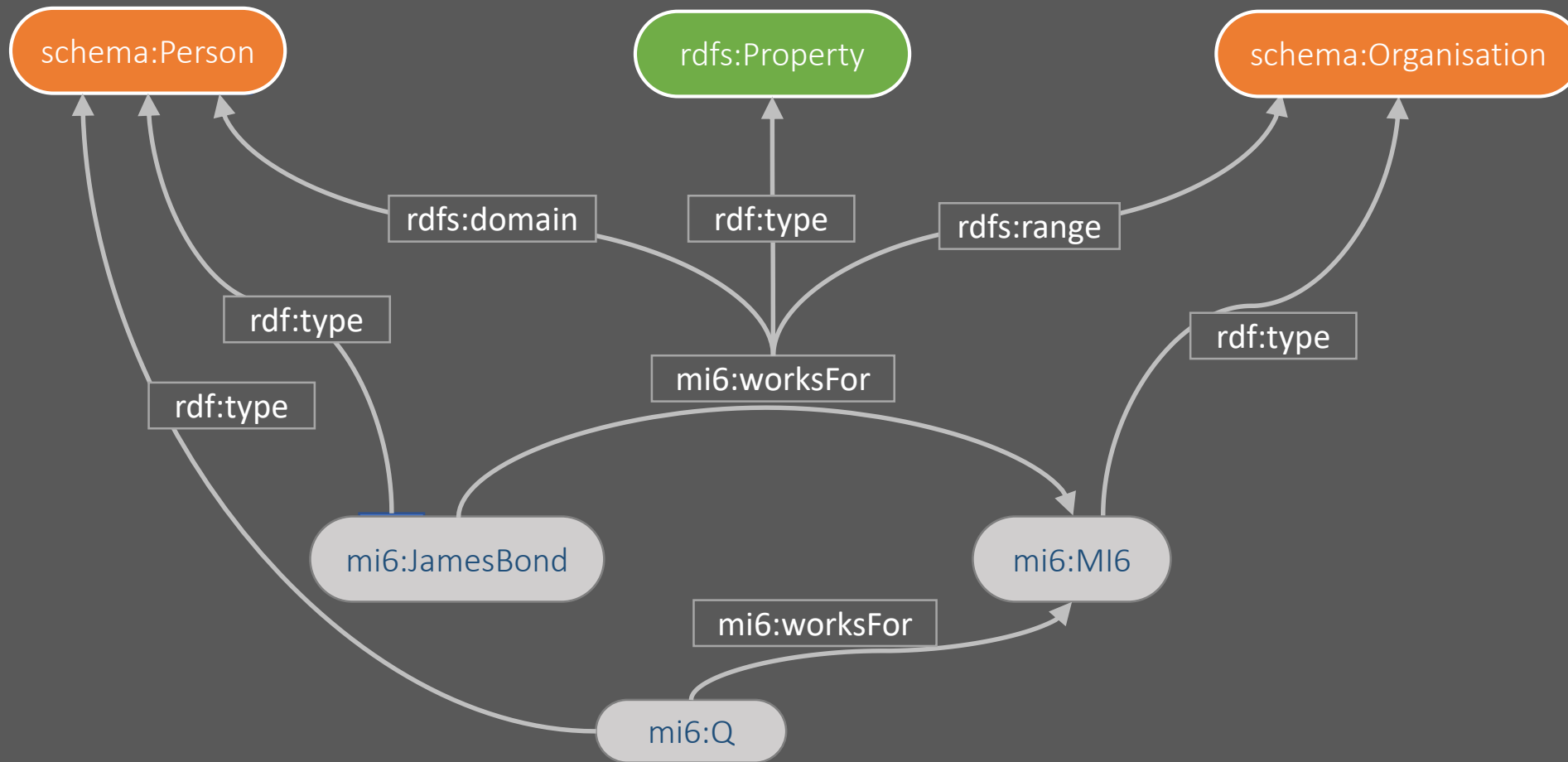
INSTANCE GRAPH

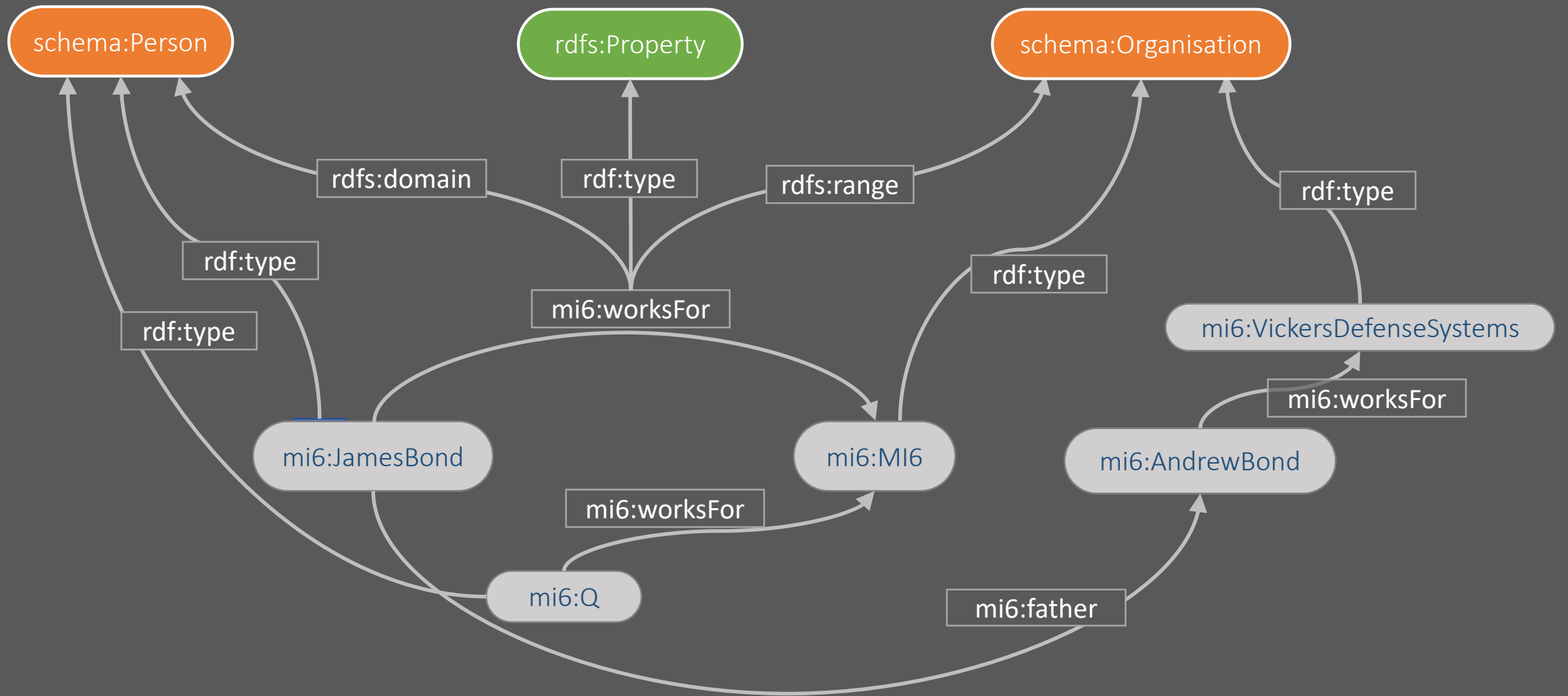


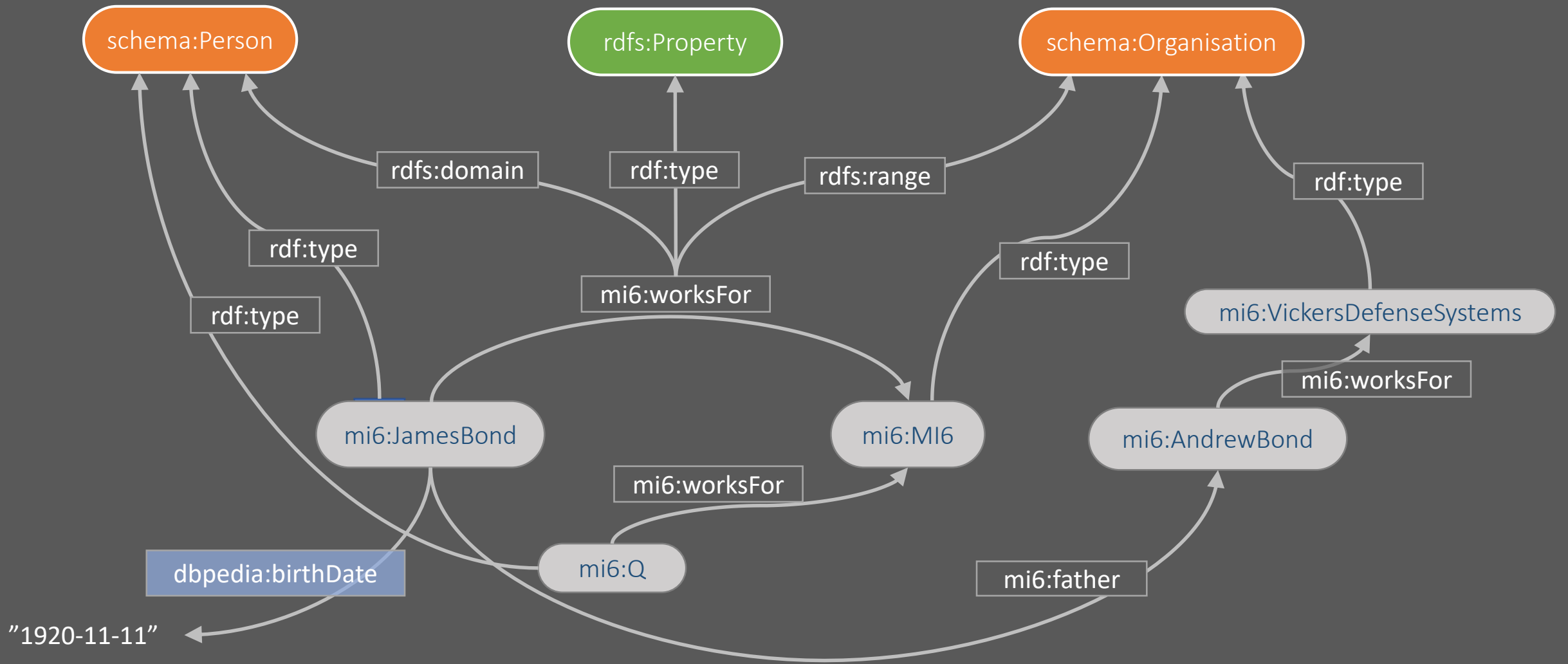
SCHEMA GRAPH

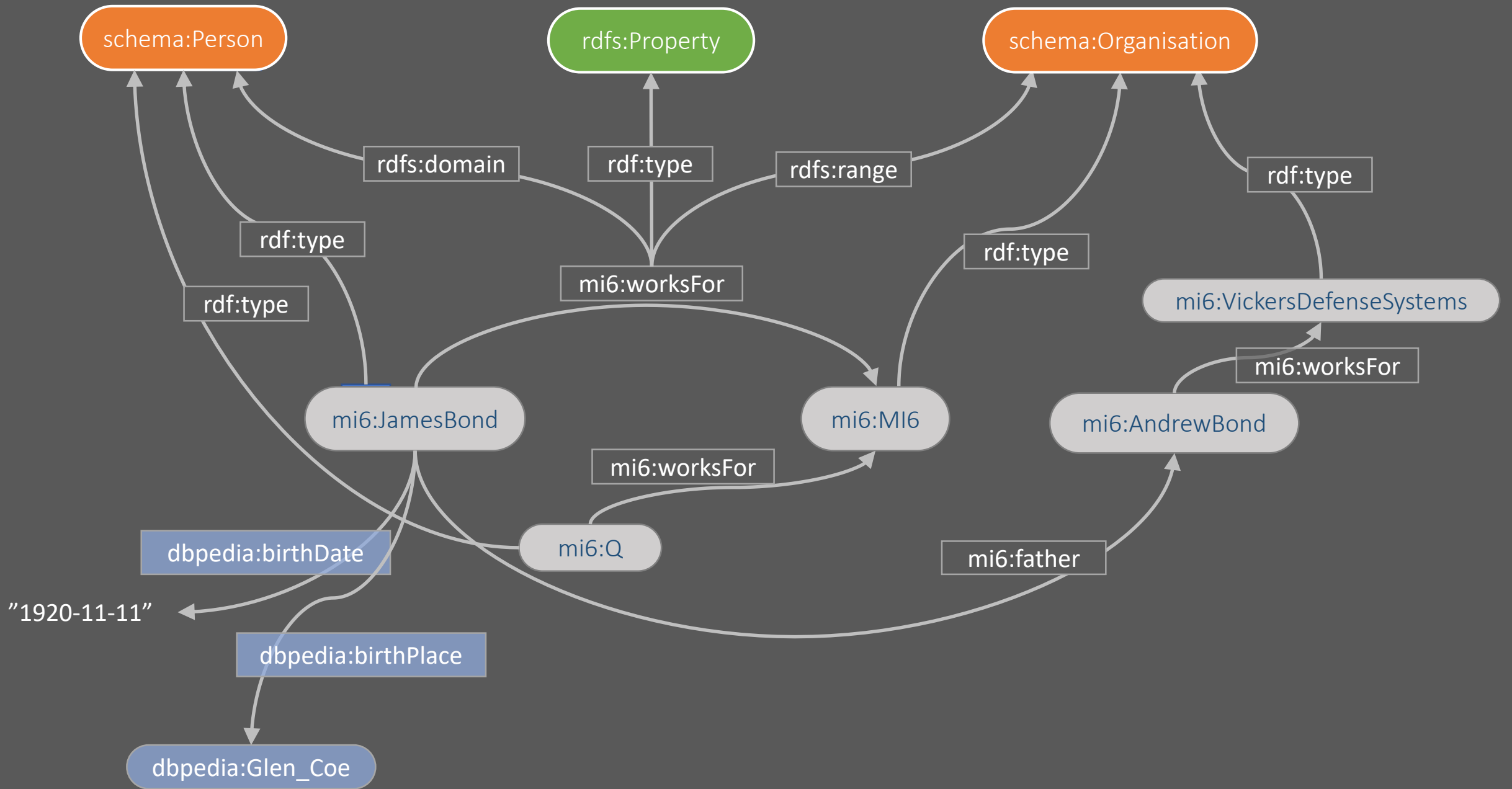


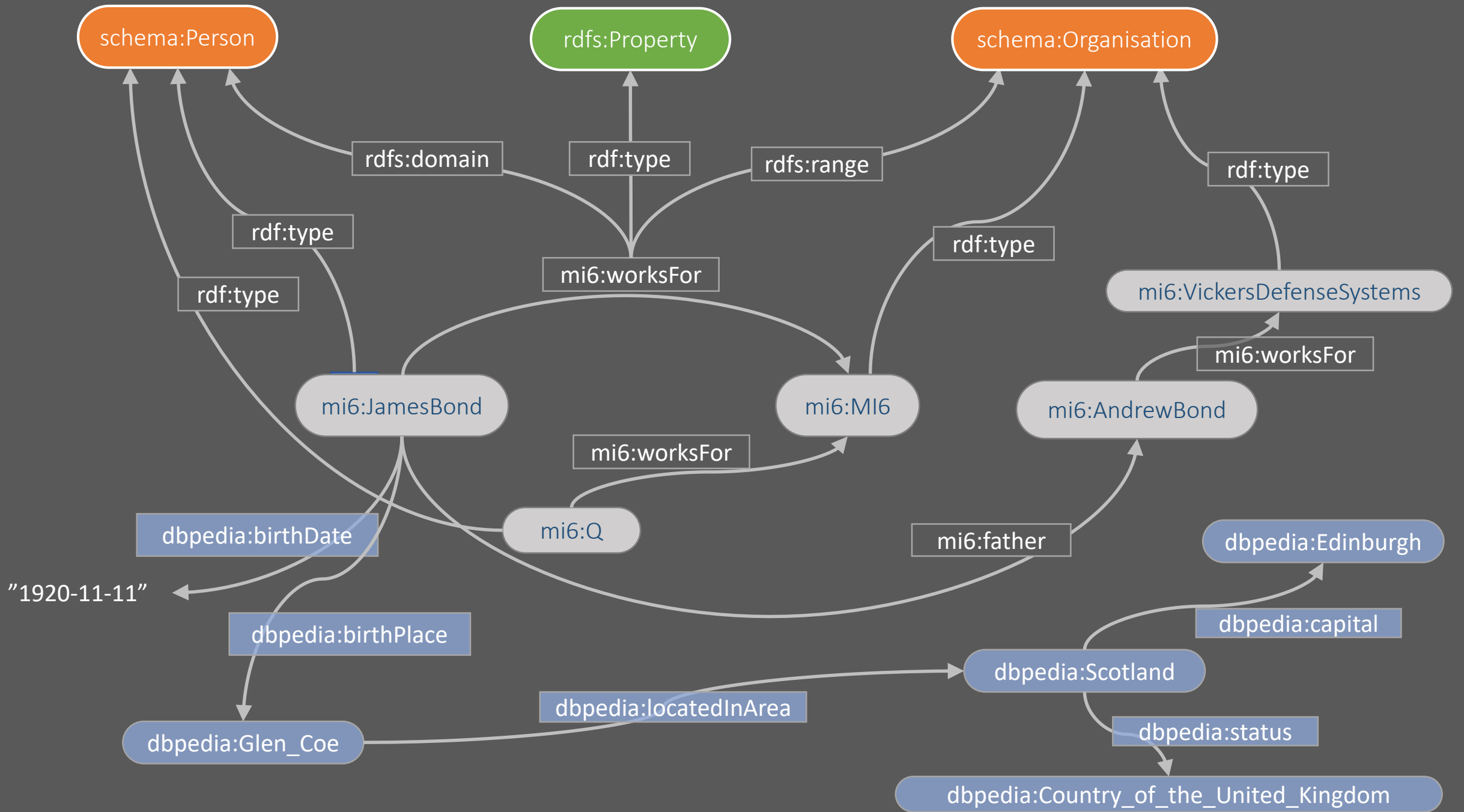
KNOWLEDGE GRAPH











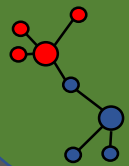
QUERYING AND INFERENCE

EXECUTE INFERENCE:

If you are born in Scotland, you are born in the UK

QUERY:

Select all persons working for MI6 that are born in the UK



LOCAL KNOWLEDGE GRAPH



DBpedia

COALESCED KNOWLEDGE GRAPH

ENTERPRISE DATA GOVERNANCE

Reporting

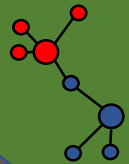
Monitoring

Compliance

Lineage

...

ENTERPRISE KNOWLEDGE GRAPH



Real time access

Model ingestion

Feature extraction

Reusable models





SYSTEM X



SYSTEM Y



GDPR

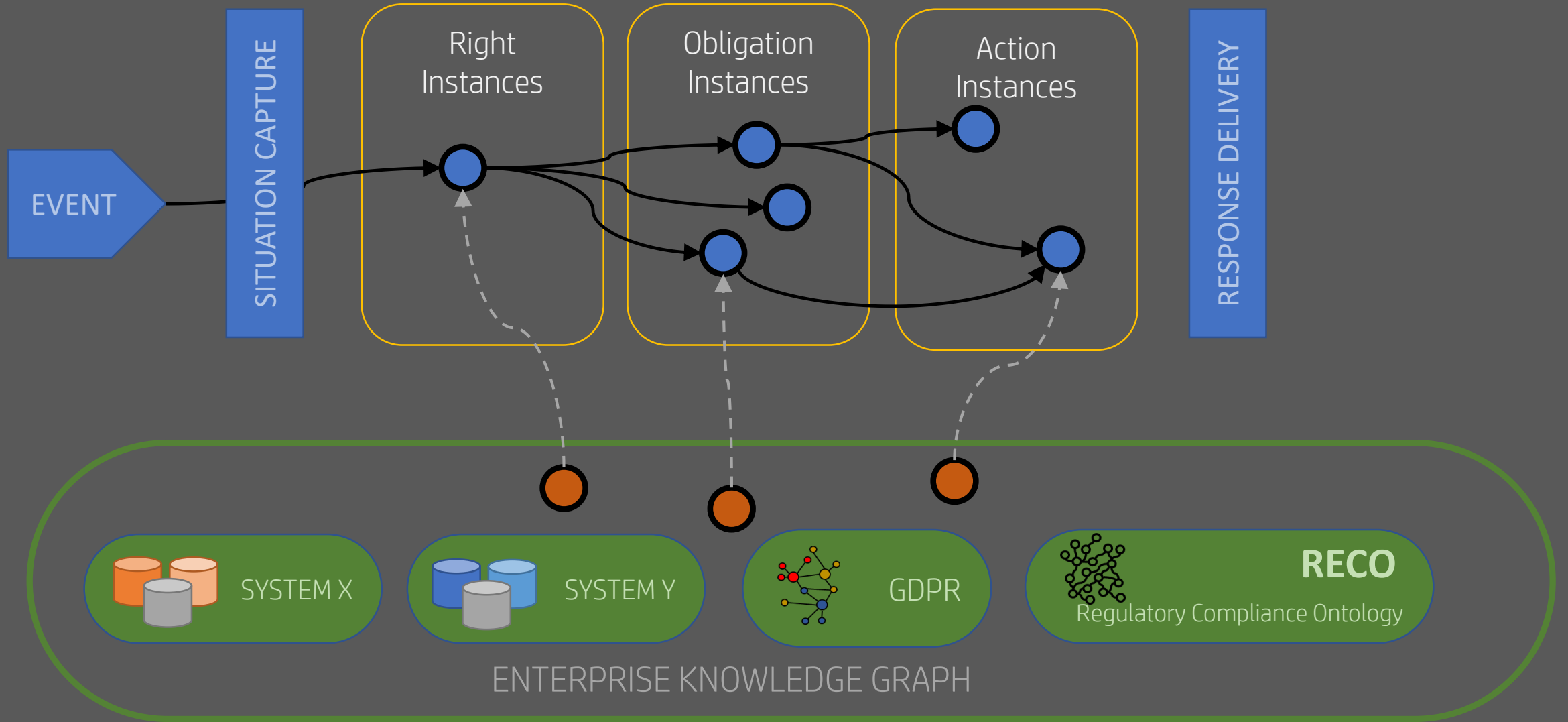


RECO

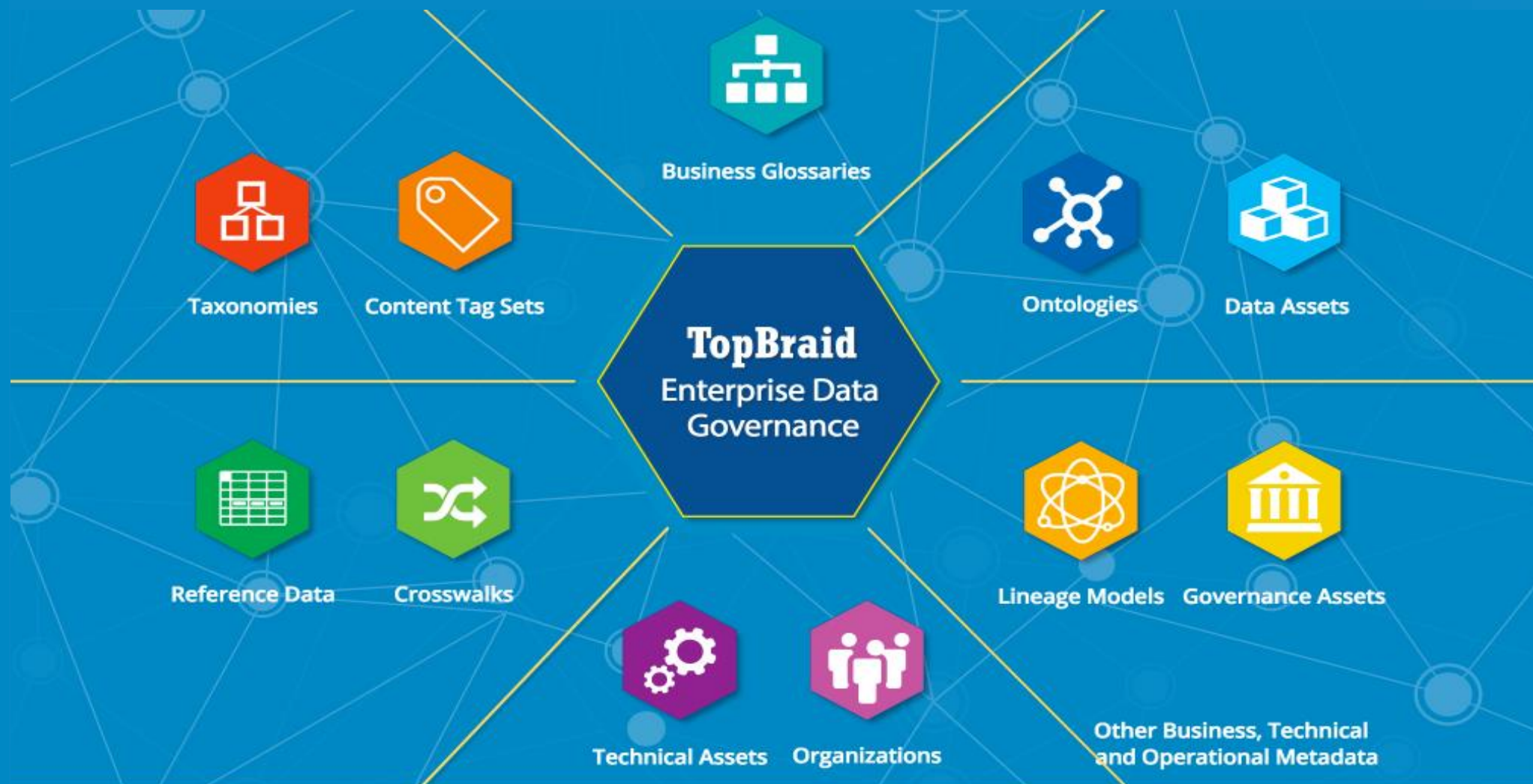
Regulatory Compliance Ontology

ENTERPRISE KNOWLEDGE GRAPH

POLICY ATTRIBUTION



Enterprise Information Management Landscape



The screenshots show the TopBraid EDG interface. The top screenshot displays a data lineage diagram with 'INPUTS', 'SOFTWARE', and 'OUTPUTS' sections. The middle screenshot shows a dashboard for 'Country Codes' with a sidebar menu and a main content area containing 'Completeness and Validity' metrics, 'Information about the Reference Dataset', and 'Semantic Analysis of Country'. The bottom screenshot shows a conceptual model diagram for 'contract' with various relationships and properties.



Collaborate

Catalog

Connect

Curate

For Governance of Enterprise Data Assets

DEMO

Knowledge Graphs & Data Governance

- > Graphs combine into graphs
- > Data and models in one representation, both are machine-readable
- > Multiple models combined in one dataset (graph)
- > Open standards

20 tables plus 20 tables
equals 40 tables

20 graphs plus 20 graphs
equals 1 graph

Knowledge Graphs & Data Governance

- > Graphs combine into graphs
- > Data and models in one representation, both are machine-readable
- > Multiple models combined in one dataset (graph)
- > Open standards

- > Amount of information
 - > Graphs are modular
 - > EDG divides the enterprise knowledge graph into intuitive, easy-to-work-with data asset collections
- > Diversity of information & Diversity of representation
 - > EDG is model-driven and can deal with data conforming to any number of models
 - > Reusability, interoperability
 - > Predefined generic models and enterprise specific models combined
 - > Models at different levels of abstraction, interlinked



Borobudur ★

4,5 ★★★★★ 1.960 Google-recensies

Tempel in Indonesië

Website

Routebeschrijving

De Borobudur is een boeddhistisch heiligdom op 40 km ten noordwesten van Jogjakarta in de provincie Midden-Java, in het centrum van het Indonesische eiland Java. [Wikipedia](#)

Adres: Indonesië

Hoogte: 35 m

Bouw gestart: 750 n.Chr.

Openingstijden: **Nu gesloten**

Architect: Gunadharna

Een bewerking voorstellen

Recensies

Een recensie schrijven

Een foto toevoegen

1.960 Google-recensies

Verzenden naar uw telefoon

Verzenden

Gerelateerd

Nog 10+ weergeven



Prambanan



Candi Mendut



Merapi



Bromo



Keraton Ngayogyakarta...
Hadiningrat

Meer over Borobudur

Feedback



F-16 Fighting Falcon

Vliegtuigmodel

De F-16 Fighting Falcon is een eenmotorige multifunctionele straaljager die sinds 1979 veel gebruikt wordt. In onder andere Nederland en België was het de opvolger van de F-104 Starfighter. In Nederland werd tevens de F-5 door de F-16 vervangen. [Wikipedia](#)

In gebruik genomen: 17 augustus 1978

Lengte: 15 m

Kosten per stuk: 14.600.000–18.800.000 USD (1998)

Eerste vlucht: 20 januari 1974

Fabrikanten: Lockheed Martin Aeronautics, Lockheed Martin, General Dynamics

Motortype: General Electric F110, Turbofan

Ontwerpers: Robert H. Widmer, Lockheed Martin

Feedback

Gezicht op Delft

Schilderij



Gezicht op Delft is een van de bekendste schilderijen van de 17e eeuwse Hollandse meester Johannes Vermeer. Het doek hangt in het Mauritshuis in Den Haag. Volgens onderzoek werd het geschilderd tussen 1659 en 1662. [Wikipedia](#)

Kunstenaar: Johannes Vermeer

Afmetingen: 98 cm x 1,18 m

Locatie: Mauritshuis

Onderwerp: Delft

Medium: Olieverf

Gemaakt: 1660–1661

Gerelateerd

Nog 15+ weergeven



Het straatje
Johannes Vermeer



Het melkmeisje
Johannes Vermeer



Meisje met de parel
Johannes Vermeer



Brieflezende vrouw in het blauw
Johannes Vermeer



De muzikles
Johannes Vermeer

Feedback

EDG is Model Driven

Core flexibility and extensibility

- Add user defined models, assets and properties as needed (model-driven)

Models: pre-built and user defined

- Support multiple types of governance assets and activities

Flexible Connections:

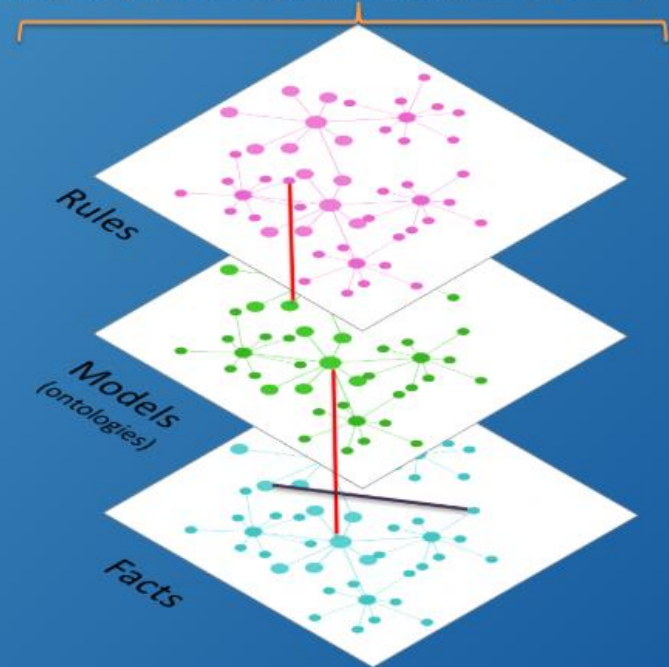
- describe relation between any type of assets or activity
- Enable people (UI) and software (APIs/web services) to view, follow and query the connections to answer core questions, e.g. “Where did this come from?”
- Bridge data “silos” for comprehensive data governance



TOPBRAID EDG – AN ENTERPRISE KNOWLEDGE GRAPH FOR DATA GOVERNANCE AND LEARNING



A trusted foundation for AI and Machine Learning





... Questions?

Thank You !



TopBraid EDG Features

- Metadata and Data Ingestion
- Discovery, Profiling and Enrichment
- Governance
- Search, Reporting and Visualization
- Collaboration
- Systems Integration
- Enterprise-ready

More on Data Governance & TopBraid EDG ...

More Webinar Recordings, Slides, Q&A: <https://www.topquadrant.com/knowledge-assets/topquadrant-webinars/>

Short Videos:

- TopBraid EDG Animated Video <http://www.topquadrant.com/EDGvideoemail/>
- TopBraid EDG Data Lineage Demo Video <http://www.topquadrant.com/knowledge-assets/videos/#edglineagedemo>
- TopBraid EDG Semantic Search Demo Video <http://www.topquadrant.com/knowledge-assets/videos/#edgsearchdemo>

Blog Posts on:

- Data Governance <http://www.topquadrant.com/category/the-semantic-ecosystems-journal/data-governance/>
- Reference Data Management <http://www.topquadrant.com/category/the-semantic-ecosystems-journal/reference-data-management/>

Data Governance White Papers

“Building Bridges to Business Context is Essential to Data Governance”

www.topquadrant.com/docs/whitepapers/Building-Bridges-to-Business-Context-is-Essential-to-Data-Governance.pdf

“How can Semantic Information Management Help You to Preserve Meaning in a Dynamic Data Environment?”

https://www.topquadrant.com/docs/whitepapers/TopQuadrant_HowCanSemantic_Whitepaper.pdf

“Maturing Reference Data Management” <http://www.topquadrant.com/solutions/maturing-reference-data-management/>

Companion white paper

“Maturing Information Governance with TopBraid EDG”

<http://www.topquadrant.com/docs/whitepapers/Maturing-Information-Governance-with-TopBraid-EDG.pdf>

- Data lineage
 - <http://www.topquadrant.com/knowledge-assets/videos/#edglineagedemo>
- Search
 - <http://www.topquadrant.com/knowledge-assets/videos/#searchdemo>
- Reference data management
 - <http://www.topquadrant.com/knowledge-assets/topquadrant-webinars/#TQ-EDG-RDM-webinar>
- Metadata management
 - <http://www.topquadrant.com/knowledge-assets/topquadrant-webinars/#TQ-EDG-metadata-mgt-webinar>