Semantics for Beginners

Rail Data Forum 2025 – Cluj-Napoca

Pierre-Antoine Champin

https://champin.net/2025/rdf/

About Pierre-Antoine

- Associate professor at Université Lyon 1, currently seconded to Inria
- Specialized in knowledge representation and exchange on the web
- Involved in several standardization groups
- Since 2021, W3C fellow

Structure of the talk

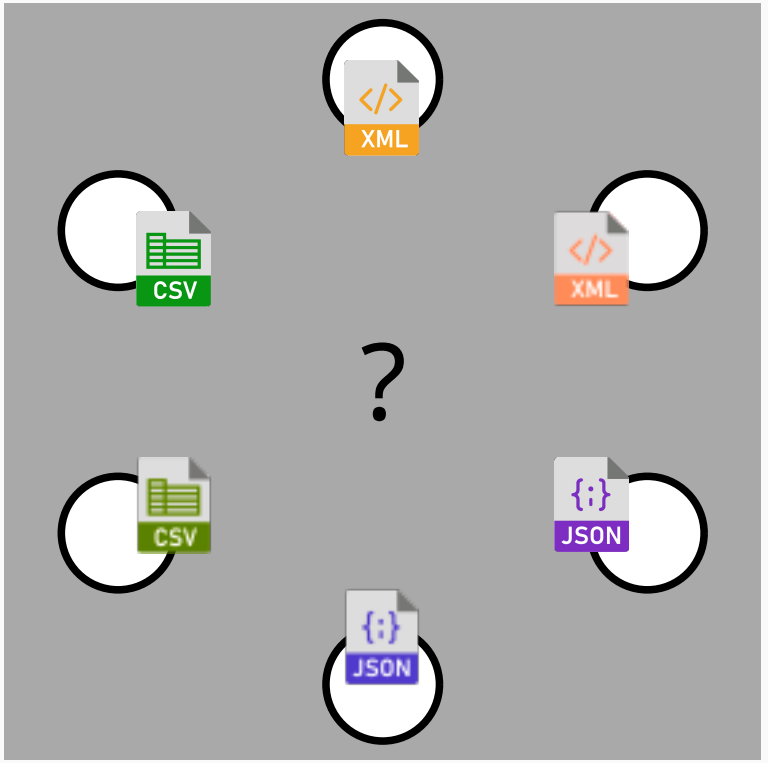
1. Semantics, what for?

2. RDF: a data model for interoperability

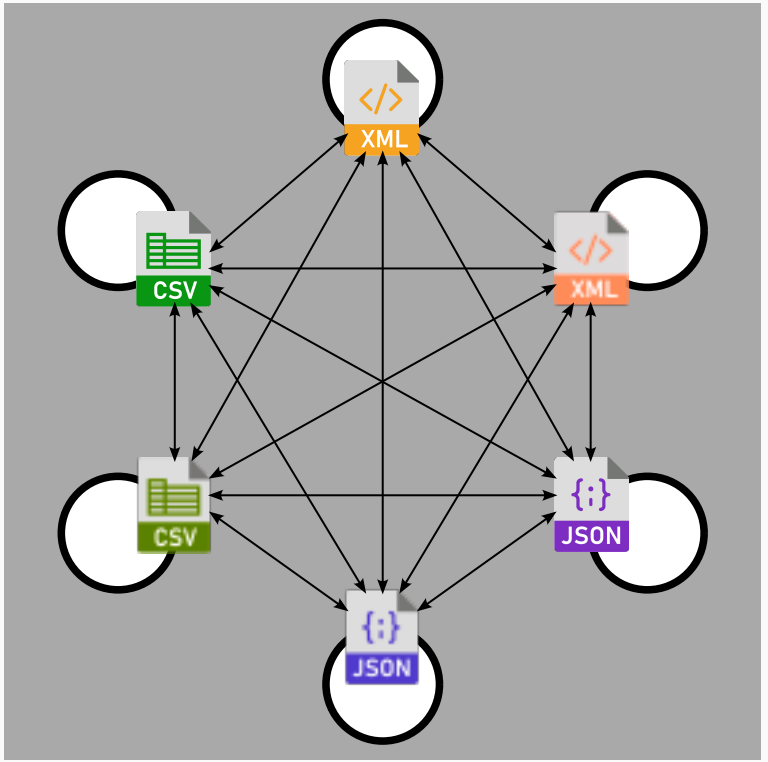
3. RDFS: a lightweight ontology language

4. OWL: expressive ontology language(s)

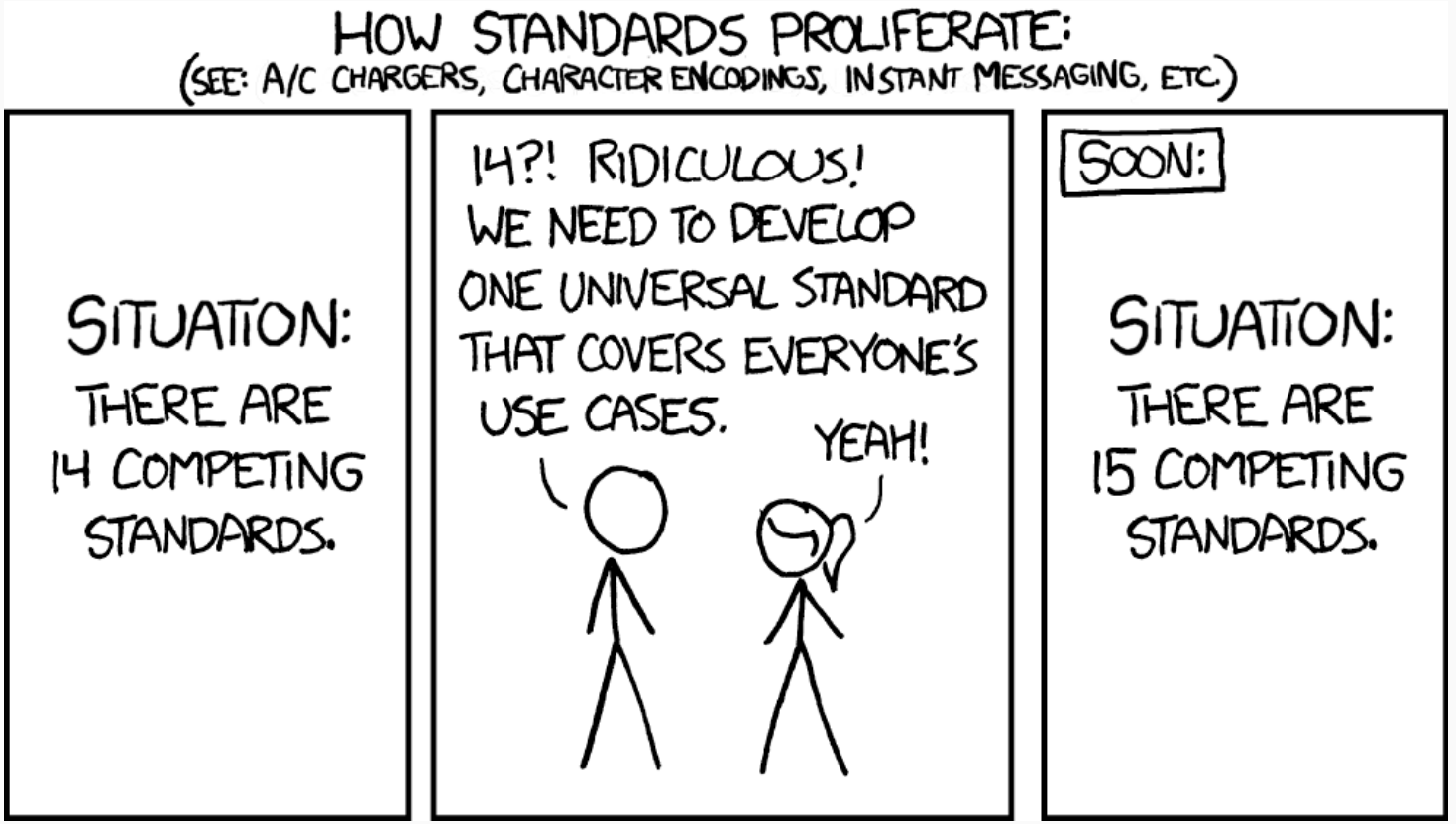
Semantics for interoperability



N^2 converters

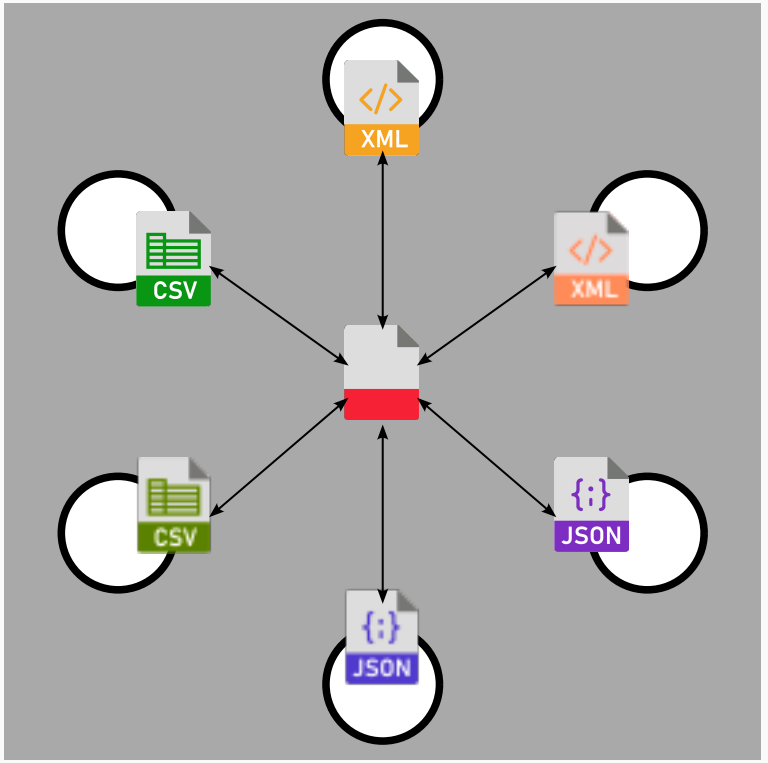


One standard to rule them all



Source image: https://xkcd.com/927/

Pivot



Two aspects of interoperability

- Syntactic:

  using the same format

- Semantic:

  using the same conceptual model

Source image: Wikimedia Commons

Syntactic / not semantic

```
<person>
  <name>Alan</name>
  <surname>Turing</surname>
  <address>
    Bletchley Park,1>
    Milton Keynes,
    UK
  </address>
</person>
```

```
<person>
  <name>Hedy Lamarr</name>
  <address>
    hedy@lamarr.name
  </address>
</person>
```

→ no real "interoperability"

Semantic / not syntactic

`name`

Name typically used to differentiate people from the same family, clan, or other social group who have a common last name.

Also known as: given name, first name

`surname`

Part of a person's name that is transmitted to a child by one of its parents.

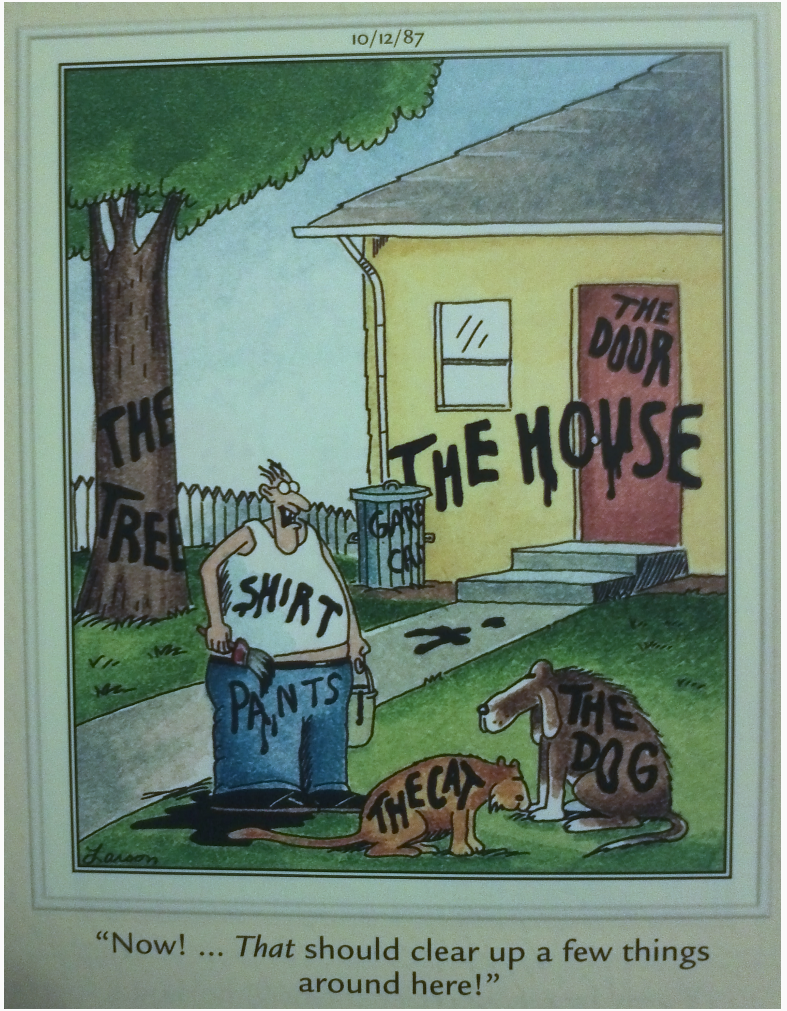
Also known as: family name, last name

...

Source definitions: Wikidata

Knowledge Engineering

- A subfield of AI
- (not *that* AI)
- Focuses on the design of ontologies, a.k.a.
  - Conceptual models
  - Conceptual schemas



Source image: The Far Side, Gary Larson

Ontology

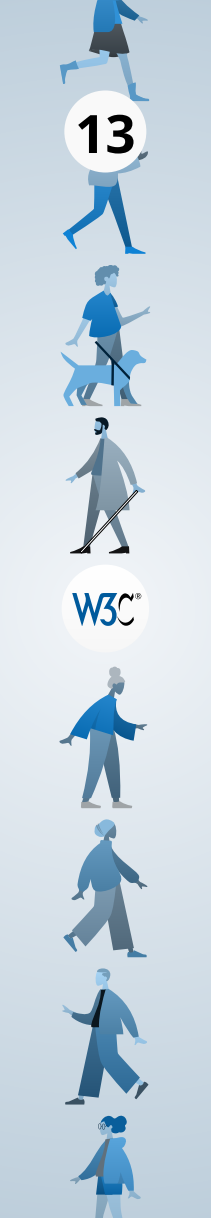
“An ontology is an explicit specification of a conceptualization”
Gruber 1993

“An ontology is a description (like a formal specification of a program) of the concepts and relationships that can formally exist for an agent or a community of agents.”
ibid.

Ontology vs. schema

- Schemas (XML, SQL, JSON...) primarily focus on syntax (the data)
- Ontologies primarily focus on semantics (the application domain)

```
<person>
  <name>Alan</name>
  <surname>Turing</surname>
  <address>
    Bletchley Park,l>
    Milton Keynes,
    UK
  </address>
</person>
```

Structure of the talk

1. Semantics, what for?

2. RDF: a data model for interoperability

3. RDFS: a lightweight ontology language

4. OWL: expressive ontology language(s)

RDF: Resource Description Frameworks

- A W3C Recommentation (1999, 2004, 2014, 202?)
- an abstract data model for achieving interoperability at Web scale
- Semantic Web, Linked Data

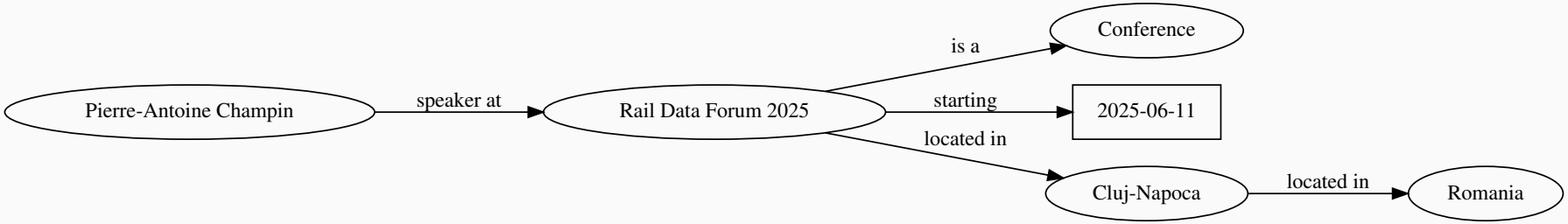
RDF triples

Information is broken down into statements (also called triples) asserting that two things are in relationship with each other. Examples:

Pierre-Antoine Champin	speaker at	Rail Data Forum 2025
Rail Data Forum 2025	is a	Conference
Rail Data Forum 2025	starting	2025-06-11
Rail Data Forum 2025	located in	Cluj-Napoca
Cluj-Napoca	located in	Romania

RDF graph

A set of RDF triples can be represented as a graph (each edge representing one statement).



Globally Unambiguous Identifiers

- Most identifiers are designed for a specific context;
- outside of this context, they becomome ambiguous
- and therefore fail to *identify*...

Examples contextual identifiers for a person

- full name
- social security number

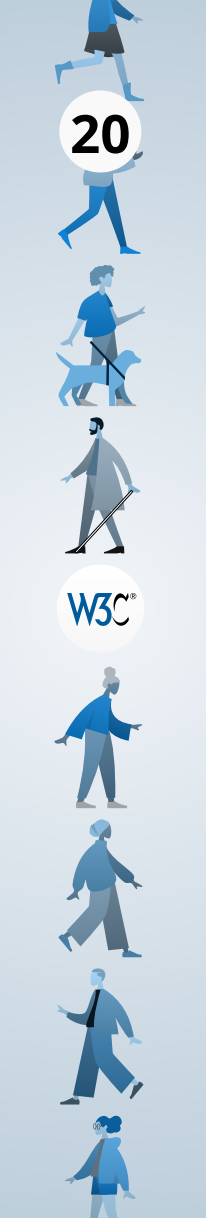
Uniform Resource Locator/ Identifier

URL, URI, IRI

- `https://champin.net/2025/rdf`
- `mailto:pierre-antoine@w3.org`
- `tel:+33-666-52-01-22`
- `doi:10.5594/SMPTE.ST2067-21.2020`
- `geo:43.615775,7.068552`
- `urn:isbn:978-951-0-18435-6`

HTTP URLs for arbitrary things

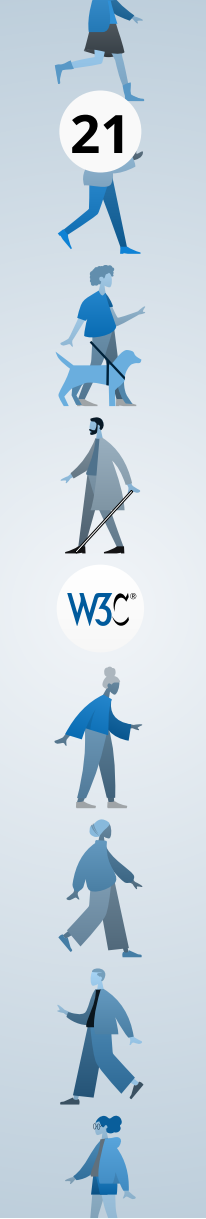
- Is it acceptable to use `http:` (or `https:`) URLs to identify persons, places, concepts?...
- Long standing controversy, but the conclusion is “yes” (with some caveats).
- (Even considered good practice for Linked Data principles)
- E.g. `https://champin.net/#pa`



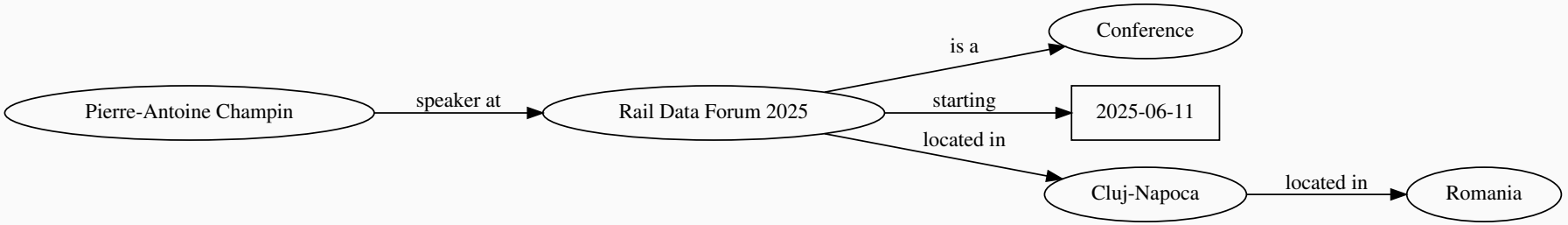
Concrete syntaxes for RDF

- RDF/XML (XML-based)
- JSON-LD (JSON-based)
- Turtle (text-based)
- N-Triples (line-based)

- interoperable (based on the abstract data model)
- domain-agnostic (only the URLs in the graph are domain-dependant)



Open World Assumption



Romania	located in	Africa ?
Romania	located in	Europe ?

Anything that is not known to be true is considered unknown (maybe true, maybe false) –until more information becomes available.

Structure of the talk

1. Semantics, what for?

2. RDF: a data model for interoperability

3. RDFS: a lightweight ontology language

4. OWL: expressive ontology language(s)

RDFS: RDF Schema

- A W3C Recommentation (2004, 2014, 202?)
- a lightweight ontology language for RDF
- in other words: an ontology for ontologies
- ontologies are themselves described in RDF

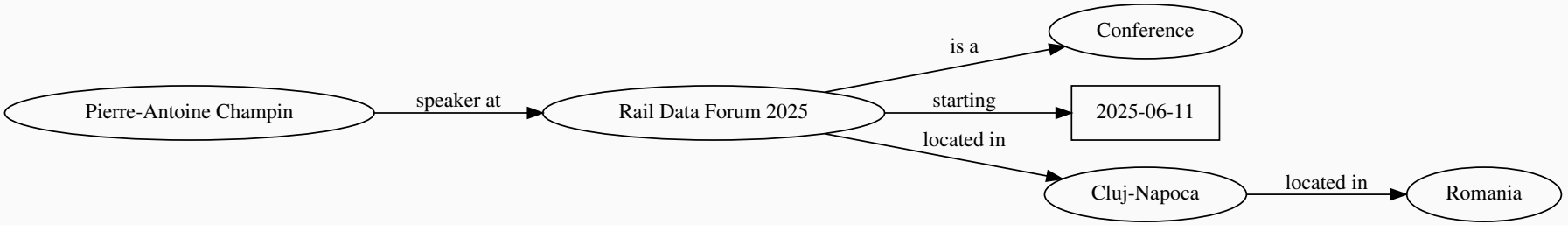
RDFS ontology

Gruber's definition:
“An ontology is a description (like a formal specification of a program) of the **concepts and relationships** that can formally exist for an agent or a community of agents.”

In RDFS, an ontology consists of

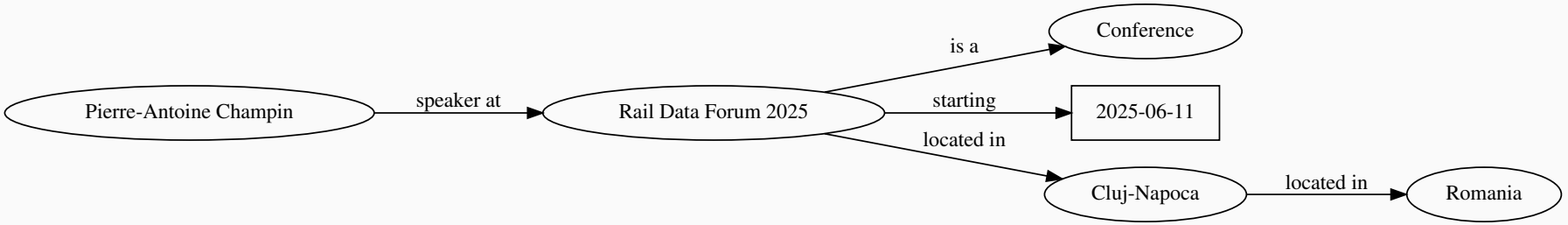
- classes (concepts)
- properties (relationships)
- special relationships between them

Classes and properties



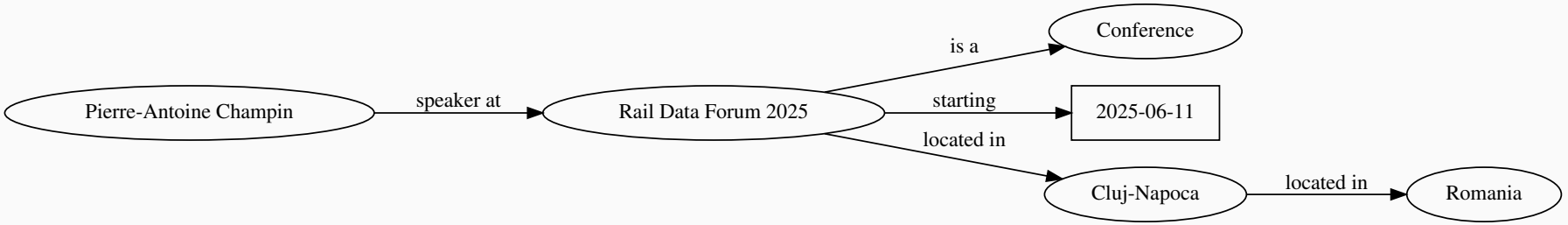
ex: speaker at	rdf: is a	rdf: **Property**
ex: Conference	rdf: is a	rdfs: **Class**
ex: starting	rdf: is a	rdf: **Property**
ex: located in	rdf: is a	rdf: **Property**
:	:	:

Label and comment



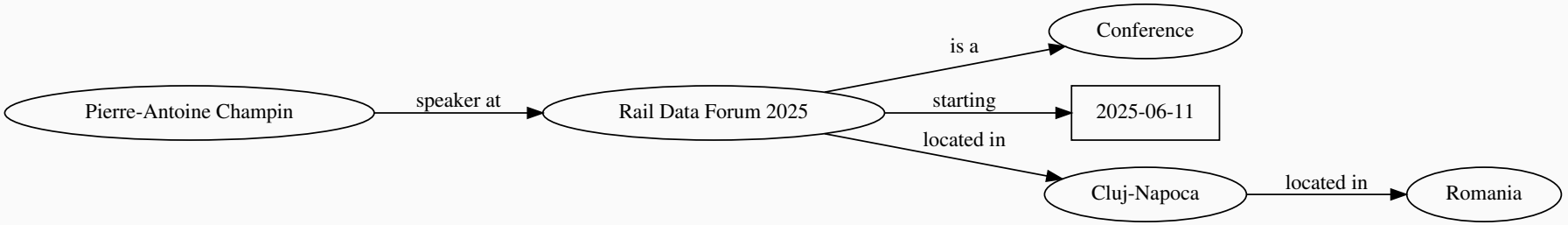
ex: speaker at	rdfs: **label**	"speaker at" @ en
ex: speaker at	rdfs: **label**	"orateur à" @ fr
ex: speaker at	rdfs: **comment**	"(detailed description)"
ex: Conference	rdfs: **label**	"conference" @ en
...	...	...

Domain and range



ex: speaker at	rdfs: **range**	ex: Conference
ex: speaker at	rdfs: **domain**	schema: Person
ex: starting	rdfs: **range**	xsd: date
ex: Location	rdf: is a	rdfs: Class
ex: located in	rdfs: **range**	ex: Location

Subclass



ex: Event

ex: Conference

ex: starting

rdf: is a

rdfs: **subclass of**

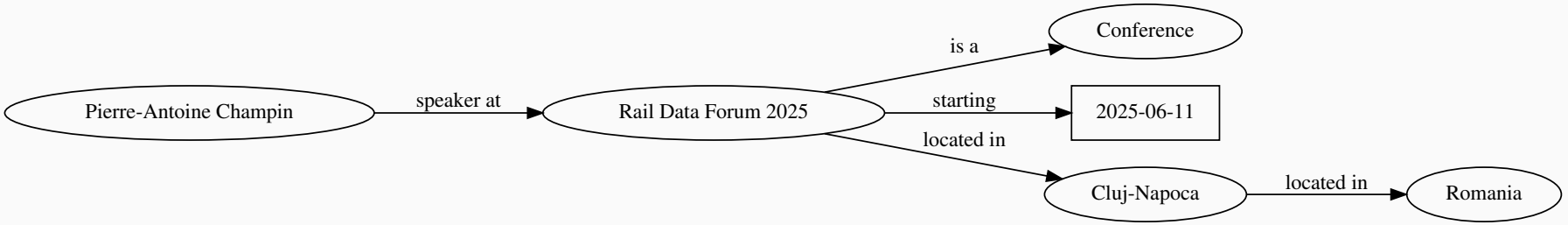
rdfs: domain

rdfs: Class

ex: Event

ex: Event

Subproperty



ex: participant in	rdf: is a	rdf: Property
ex: participant in	rdfs: range	ex: Event
ex: participant in	rdfs: domain	schema: Person
ex: speaker at	rdfs: **subproperty of**	ex: participant in

Structure of the talk

1. Semantics, what for?

2. RDF: a data model for interoperability

3. RDFS: a lightweight ontology language

4. OWL: expressive ontology language(s)

OWL Web Ontology Language

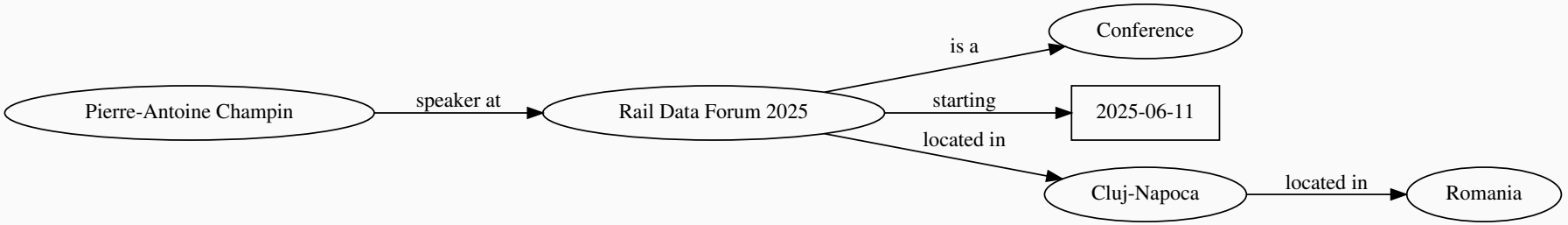
- A W3C Recommentation (2004, 2012)
- a powerful (family of) language(s) for expressing ontologies

Source image: Wikimedia Commons / Disney Fandom

OWL Web Ontology Language

- OWL ontologies are also defined in terms of classes and properties
- although OWL refines these notions compared to RDFS
- OWL reuses many terms from RDFS: label, comment, domain, range, subclass of, subproperty

Forms of negation



ex: Event

owl: **disjointWith**

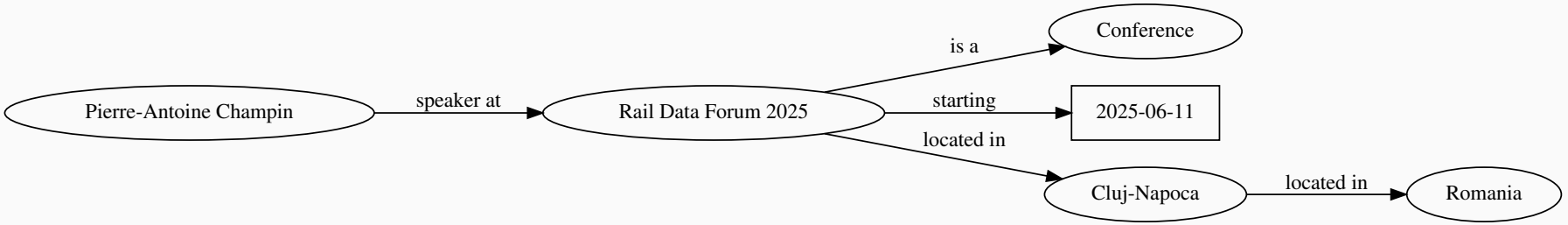
schema: Person

ex: located in

owl: **propertyDisjointWith**

ex: next to

Specifying properties



ex: located in

rdfs: is a

ex: **TransitiveProperty**

ex: next to

rdf: is a

owl: **SymmetricProperty**

ex: starting

rdf: is a

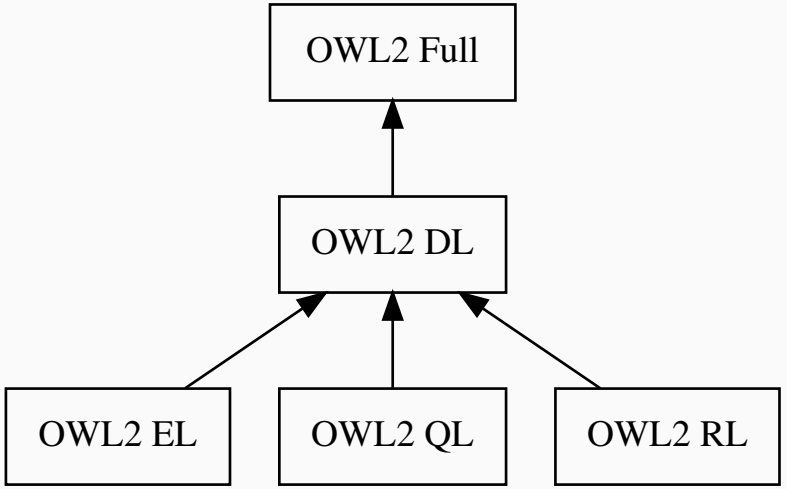
owl: **FunctionalProperty**

but also reflexive, irreflexive, asymmetric, inverse-functional, inverse, property path...

Specifying classes

- OWL allows to define classes using classical set operations (union, intersection, complement)
- e.g. the class of all Events that are not Conferences ($E \cap C^c$)
- e.g. the class of all Men and Women ($M \cup W$)
- OWL allows to define classes based on the properties of their instances
- e.g. the class of all persons who participate in at least 2 conferences
- e.g. the class of all events that are located in Romania

OWL Profiles



See https://www.w3.org/TR/owl2-profiles/

Any question?

https://champin.net/2025/rdf/