U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

Semantic Web and Linked Data

M.EIC FEUP

Liliana Ferreira

S1 2025/26

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

Class 3: Learning Objectives

- Resource Description Framework:
  - Understand RDF principles and how to provide useful RDF information
  - RDF vocabulary and serialization.
- Practical Sheet 1.
- Practical Work Development.

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

Key Concepts

- What is Web Semantics?
  - Turning the web into a database through structured, meaningful data.

- What is Linked Data?
  - Connecting data across different domains for more meaningful, interoperable use.

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

Web Semantics

- Semantic Web: Giving data meaning (metadata)

- Key Technologies:
  - RDF (Resource Description Framework)
  - OWL (Web Ontology Language)
  - SPARQL (Query Language)

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

How does a machine interpret this:

“ Ronaldo was born in Madeira.
He plays football for Portugal.”

The Web today is still…

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

Currently most of the Web content is suitable for human use.

Typical uses of the Web today are information seeking, publishing, and using, searching for people and products, shopping, reviewing catalogues, etc.

Dynamic pages generated based on information from databases but without original information structure found in databases.

Limitations of the Web Search today

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

The Web search results are high recall, low precision.

Results are highly sensitive to vocabulary.

Results are single Web pages.

Most of the publishing contents are not structured to allow logical reasoning and query answering.

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

The Goal:

Machine Accessible Content

U.PORTO
FEUP FACULDADE DE ENGENHARIA UNIVERSIDADE DO PORTO

Resource Description Framework (RDF)

- RDF is a language for the representation of resources: **a resource can be anything**;
- A standard of W3C;
- RDF is a data model;
- One of main applications: data integration.
  - Relationships *between* documents;
- Basic building block: **triples** or statements

```
<subject, property, object>
<"Mozart", composed, "The Magic Flute">
```

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

RDF is…

- …simple: everything is triples
- …a data model: it is **not** a file format
- …a logical formalism: it has **formal semantics**
- …more than XML or JSON: XML and JSON have a *tree-based* model, RDF has a *graph-based* model
- …a Web standard: a W3C recommendation
- …the *lingua franca* for the Semantic Web

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

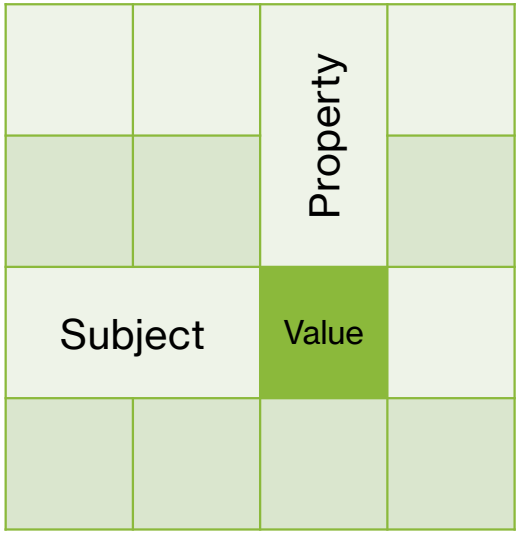
RDF means

- Resource:
  - Everything that can have an URI/IRI: pages, chairs, persons, pens, ideas
- Description:
  - Attributes, characteristics and relations between resources
- Framework:
  - Model, language and syntaxes for these descriptions

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

Resource Description Framework (RDF)

- Universal, machine-readable exchange format;
- Data structured in graphs (vertices, edges).
- Any relational data can be represented as triples:
  - Triples are statements about things (resources), using URIs or literal values



U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

RDF basics

- The core features of RDF are:
  - **Identify** things (resources)
  - Express **relations** between things (properties)
- Additional important features are:
  - Assign **data values** to things (literals)
  - Organize things in **categories** (i.e., classes or types)
  - Add **simple knowledge** about categories and relations

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

RDF decomposes descriptions into triples

(subject, predicate, object)

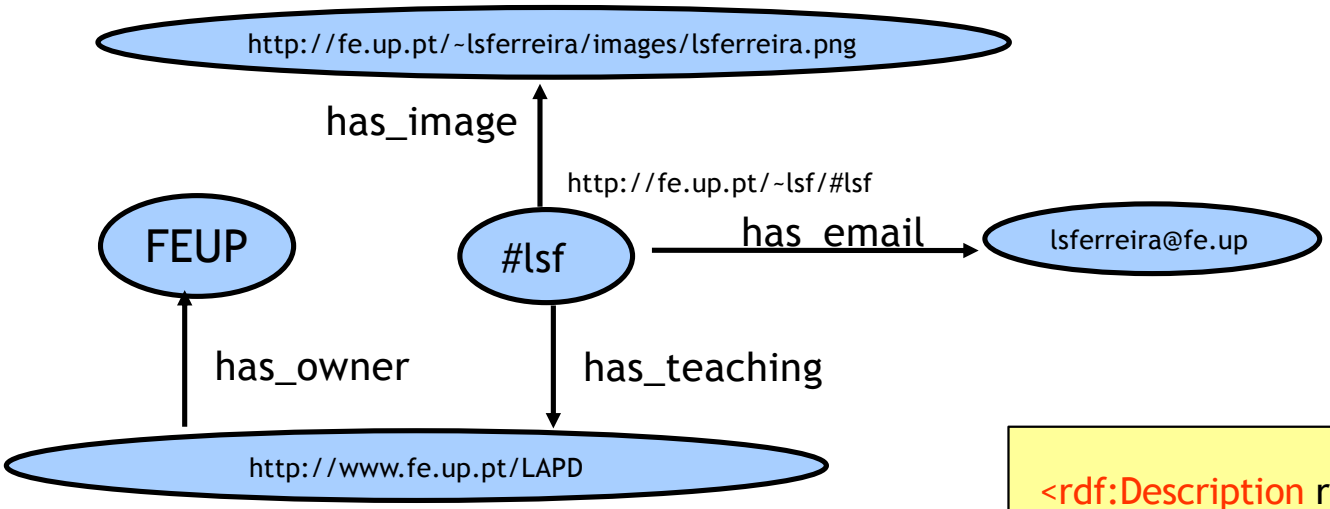
E.g.: “WSDL” has as students João, Sara and Miguel and has as topic Semantic.

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

Composition Rules for RDF Triples

1. The subject is always a resource (and not a literal)
2. The type of the binary property is identified by a URI
3. The value is a resource or a literal

RDF triples form graphs




```
<rdf:Description rdf:about="#lsf">
   <has_email>lsferreira@fe.up</has_email>
</rdf:Description>
```

U.PORTO
FEUP FACULDADE DE ENGENHARIA UNIVERSIDADE DO PORTO

Identify things

Resources

- RDF is used to **describe resources**
- A ***resource*** can be anything (real or imaginary entity, abstract or concrete)
- Every resource has a URI (Universal Resource Identifier) or IRI (Internationalized Resource Identifier)
- A URI/IRI **identify** things but *may* be used as **locators** (URL, a web address) at the same time;
- An identifier does not necessarily enable access to a resource;
- To describe a resource, it must be named or identified
- On the Web, the identification mechanism must be uniform at Web scale: an identifier must identify the same thing **everywhere** on the Web

Examples:

- ✓ `urn:ietf:rfc:3987`
- ✓ `svn://yadiyada.foo.bar/`
- ✓ `mailto:lsferreira@fe.up.pt`
- ✓ `ftp://ftp.up.pt/#meta`
- ✓ `http://en.wikipedia.org/wiki/User:Wikiuser100`

Note: to shorten notations, we use namespace **prefixes**, e.g.,
`rdf:` is for `http://www.w3.org/1999/02/22-rdf-syntax-ns#`

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

How to choose an IRI?

- If possible, reuse an existing IRI from an authoritative source, e.g.:
  - from a national library
  - from a government website, a ministry
- If not, make your own IRI:
  - use HTTP IRIs
  - use a namespace under your control
  - Cool URIs don’t change
  - Refer to the guide on Cool URIs for the Semantic Web

U.PORTO
FEUP FACULDADE DE ENGENHARIA UNIVERSIDADE DO PORTO

Relate things

Predicate

- A ***predicate*** is a specific aspect of a resource.
- It can be a characteristic that belongs to a resource, or a relationship that links one resource with another: “João **studies** at FEUP”
- Predicates describe relations between resources;
  - For example: “written by”, “composed by”, “title”, “topic”, etc;
- The predicate in RDF is also identified by IRIs. This provides a global, unique naming scheme.

Example:

```
http://example.org/data/Joao , subject
http://social.relations.com/knows ,predicate
http://example.org/data/Francisco , object
```

U.PORTO
FEUP FACULDADE DE ENGENHARIA UNIVERSIDADE DO PORTO

Data values

- As everything else, a data value (number, string, date) is a resource
- A specific data value can be identified with a **literal**, a character string that represents the value
- Every literal is typed such that its string representation can be interpreted as the correct value
- Usually, we use standard data type IRIs from the xsd: namespace (XML Schema Datatypes) and the rdf: namespace

Example:

- E.g., "42" represents the number fourty two if this is of type decimal integer, but represents sixty six if it is an hexadecimal integer

U.PORTO
FEUP FACULDADE DE ENGENHARIA UNIVERSIDADE DO PORTO

Statements

- A ***statement*** is a piece of description about a particular resource in the RDF format: an object-attribute-value triple;
- It consists of a resources, a property, and a value (subject, predicate, object)

http://mitpress.mit.edu/catalog/item/default.asp?ttype=2&tid=10140

U.PORTO
FEUP FACULDADE DE ENGENHARIA UNIVERSIDADE DO PORTO

Resource Description Framework (RDF)

- A statement about a resource instance has:
  - the resource's identifier
  - one of the resource's property (defined in an RDF schema)
  - the value for that property (can be either a literal, or a resource)

```
<rdf:RDF
  xmlns:wc="http://www.lapd.fe.up.pt/~exRDF/wc/schema">
  <rdf:Description about="http://www.cnn.com/2000/HEALTH/cancer/12/06/
            colon.cancer.ap/index.html">
  <wc:Title>Cigarette smoking linked to colorectal cancer </wc:Title>
  </rdf:Description>
</rdf:RDF>
```

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

RDF is an oriented labeled multigraph model

- RDF is an oriented labeled multigraph model
  1. Several edges can connect the same two nodes;
  2. Edges are oriented: the head is the object, the tail is the subject;
  3. Edges and nodes are labeled.

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

RDF Syntax

- The RDF data model provides an abstract, conceptual framework for defining and using metadata.
- A concrete syntax is also needed for the purposes of creating and exchanging this metadata.

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

RDF Vocabulary

- RDF defines a number of resources and properties;
- RDF vocabulary is defined in the namespace:
  - http://www.w3.org/1999/02/22-rdf-syntax-ns#

The vocabulary defined by the RDF specification is as follows:

- Classes:
  - `rdf:Property, rdf:Statement,rdf:XMLLiteral`
  - `rdf:Seq,rdf:Bag,rdf:Alt,rdf:List`

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

RDF Vocabulary

- Properties:
  - `rdf:type,rdf:subject,rdf:predicate,rdf:object,`
  - `rdf:first,rdf:rest,rdf:_n`
  - `rdf:value`

- Resources:
  - `rdf:nil`

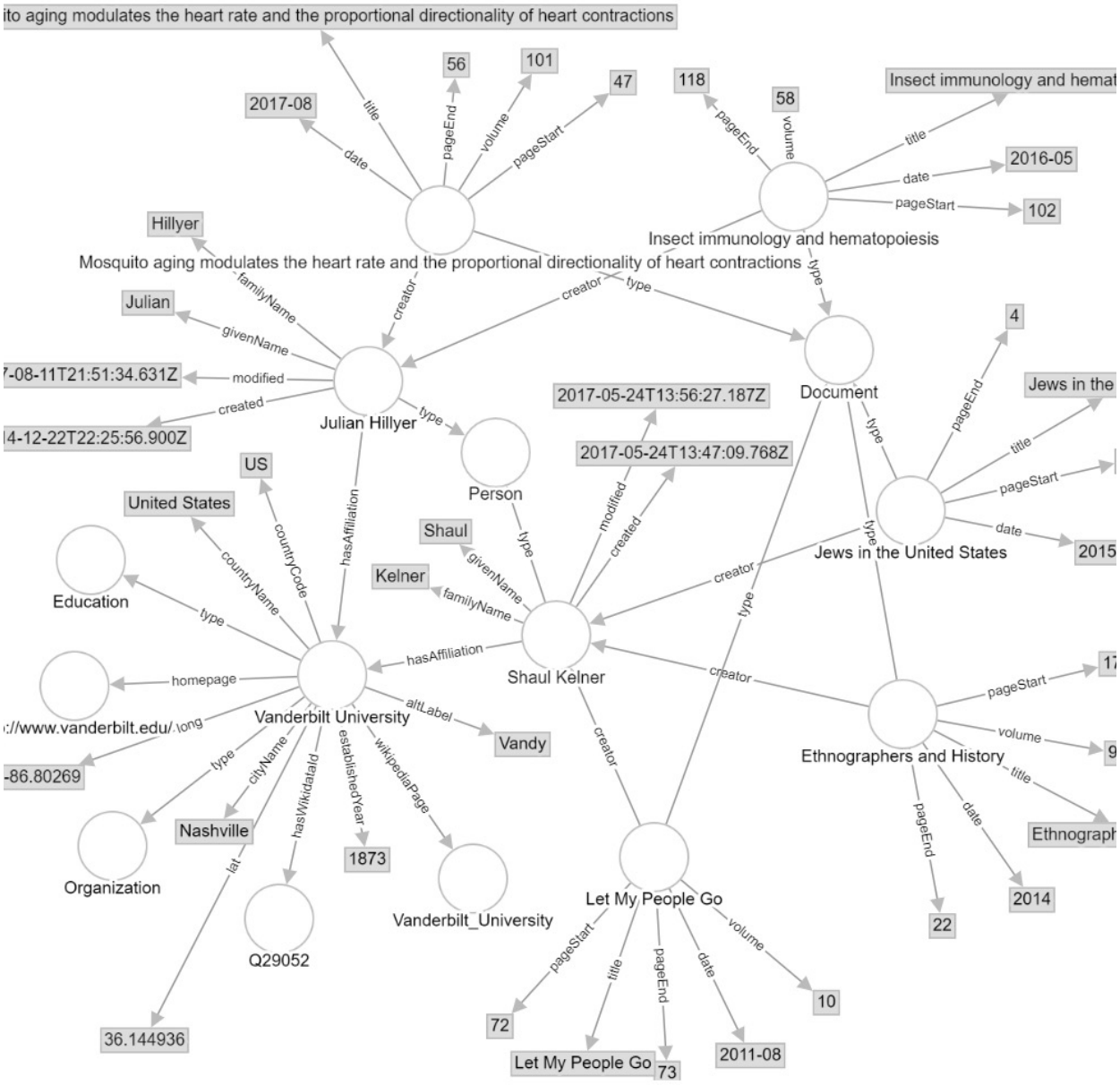
U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

RDF Serializations and Triplestores

- Since RDF is an abstract model for expressing information about graphs, it can be expressed in a number of concrete ways.
- One way that is particularly easy for humans to understand is a graphical diagram.

RDF Serializations

- The triples in this table form a graph that can be represented by this diagram.



U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

RDF Serializations

- However, it is generally not possible for machines to interpret graphs that are expressed as diagrams.
- Machines need an RDF *serialization*, a method of transmitting or storing the information about the triples in the graph as a file.

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

RDF graphs as files

- In WSLD, we will mostly use Turtle
- Others:
  - There is an XML-based syntax: RDF/XML
  - There is a JSON-based syntax: JSON-LD
  - There is an easy to parse, line-based triple syntax: N-Triples
  - There is a syntax to embed RDF in HTML and XML documents: RDFa

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

The Turtle RDF syntax

- Turtle stands for "**Terse RDF Triple Language".**
- N-Triples is a subset of the RDF Turtle serialization, meaning that any file that is valid N-Triples is also valid Turtle serialization.
- However, Turtle allows compact URIs (CURIEs) and also allows shortcuts to prevent repeating parts of triples.

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

The Turtle RDF syntax

- For example, if several triples share the same subject, the predicates and objects can be listed, separated by semicolons.

U.PORTO
FEUP FACULDADE DE ENGENHARIA UNIVERSIDADE DO PORTO

The Turtle syntax

- Full IRIs:

```
<http://www.example.com/test#this>
```

- A simple triple:

```
<http://www.example.com/test#this>
        <http://relations.example.com/in>
                <http://www.example.com/test#box> .
```

- Abbreviated IRIs (declare prefixes at the beginning of the file):

```
# This is a comment
@prefix ex: <http://www.example.com/test#> . # end dot!
@prefix rel: <http://relations.example.com/> .
ex:this rel:in ex:box . # Another comment
```

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

The Turtle RDF syntax

- The namespace prefixes that are used in the triples must be listed in a prolog at the start of the document.
- Notice that URIs aren’t required to be abbreviated.

U.PORTO
FEUP FACULDADE DE ENGENHARIA UNIVERSIDADE DO PORTO

The Turtle syntax

- Literals:

```
ex:this rel:date "2019-09-13"^^xsd:date . # normal literal
ex:this rel:name "this"@en . # language-tagged literal
ex:this rel:code "TX32" . # xsd:string can be omitted
ex:this rel:number 42 . # xsd:integer (no quotes)
ex:this rel:sizeInMeters 3.75 . # xsd:decimal (use a dot)
```

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

The Turtle syntax

- If two triples share both the same subject and predicate, the two objects can be separated by commas. For example:

```
ex:box rel:contains ex:this .
ex:box rel:contains ex:that .
# can be written
ex:box rel:contains ex:this, ex:that . # comma
```

U.PORTO
FEUP FACULDADE DE ENGENHARIA UNIVERSIDADE DO PORTO

The Turtle syntax

- Repeat object

```
ex:this rel:date "2019-09-13"^^xsd:date;
rel:name "this"@en; # new lines are optional
rel:code "TX32";
rel:nextTo ex:that, ex:thoot, ex:thus .
```

U.PORTO
FEUP FACULDADE DE ENGENHARIA UNIVERSIDADE DO PORTO

The Turtle syntax

- Turtle also allows a special abbreviation for the important predicate `rdf:type`. It can be replaced with `a` .
- Hence, the triple:

```
<http://dbpedia.org/resource/Bob_Marley> <http://www.w3.org/1999/02
/22-rdf-syntax-ns#type> <http://xmlns.com/foaf/0.1/Person> .
```

can be shortened in Turtle to:

```
dbr:Bob_Marley a foaf:Person
```

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

The Turtle syntax

- RDF text files in Turtle serialization are usually given the file extension .ttl

- Let us learn to read and write Turtle in an online editor. Go to: https://perfectkb.github.io/yate/

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

Further reading

- Semantic Web Stack
- RDF 1.1 Primer
- RDF 1.1 Concepts and Abstract
- https://www.w3.org/TR/turtle/

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

Wrap-Up and Key Takeaways

- RDF gives **structure and meaning** to web data.
- Every entity and relation is identified by a **URI**.
- RDF forms **graphs** — not trees.
- It's the foundation for **querying (SPARQL)** and **reasoning (OWL)**.
- RDF enables **interoperability**, crucial for data integration (e.g., manufacturing diagnostics, biomedical KGs, etc.).

U.PORTO
FEUP FACULDADE DE ENGENHARIA
UNIVERSIDADE DO PORTO

Class 3: Learning Objectives

- Real-World Applications & Group Exercise: Wrap-up and key takeaways
- Resource Description Framework:
  - Understand RDF principles and how to provide useful RDF information
  - RDF vocabulary and serialization.

- **Practical Sheet 1**
- **Practical Work Development**