Institut Mines-Télécom

cdt:ucum: describe measurements with The Unified Code for Units of Measure.

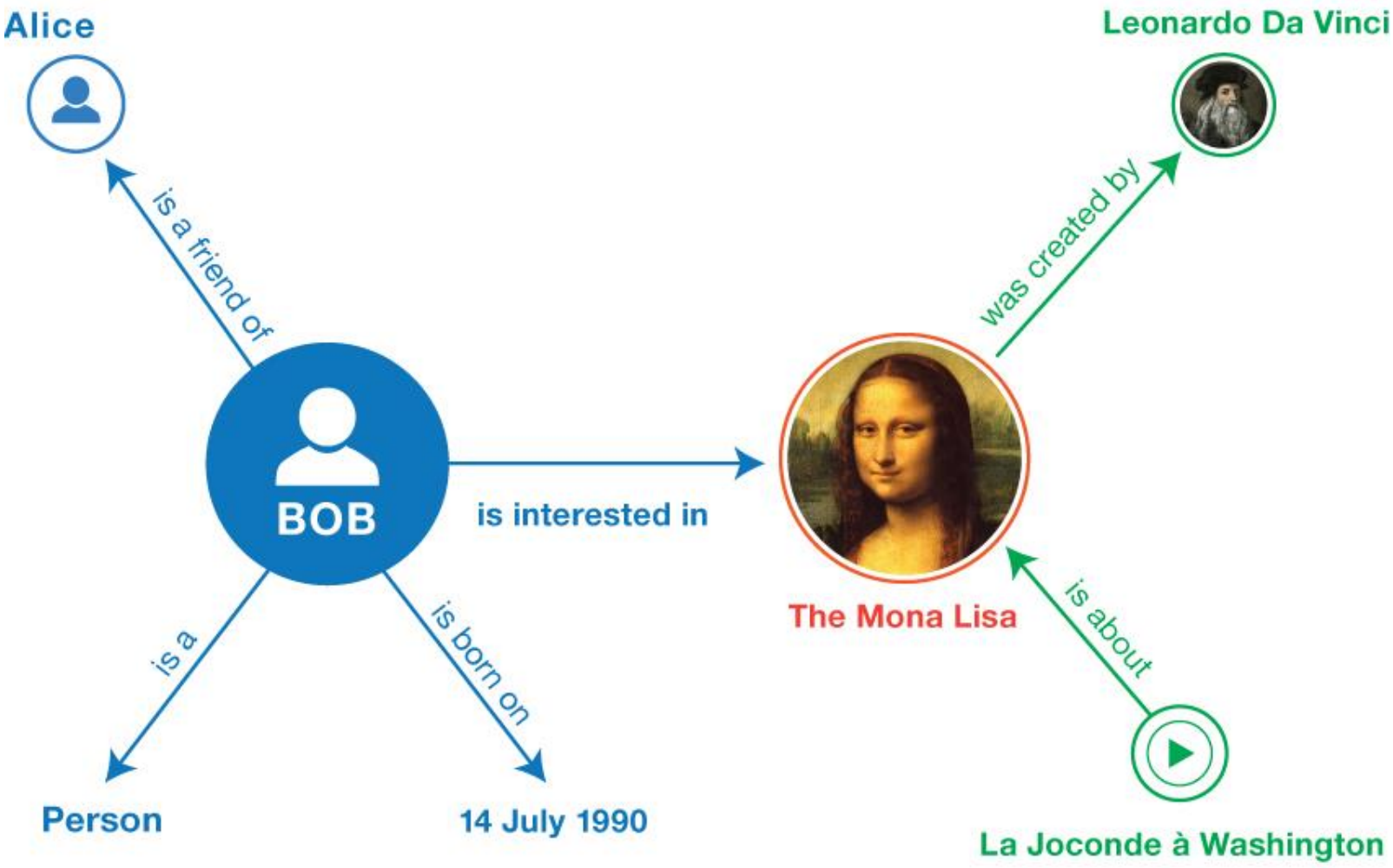
Maxime Lefrançois
http://maxime-lefrancois.info/

MINES Saint-Étienne – Institut Henri Fayol
Laboratoire Hubert Curien UMR CNRS 5516

MINES Saint-Étienne
Une école de l'IMT

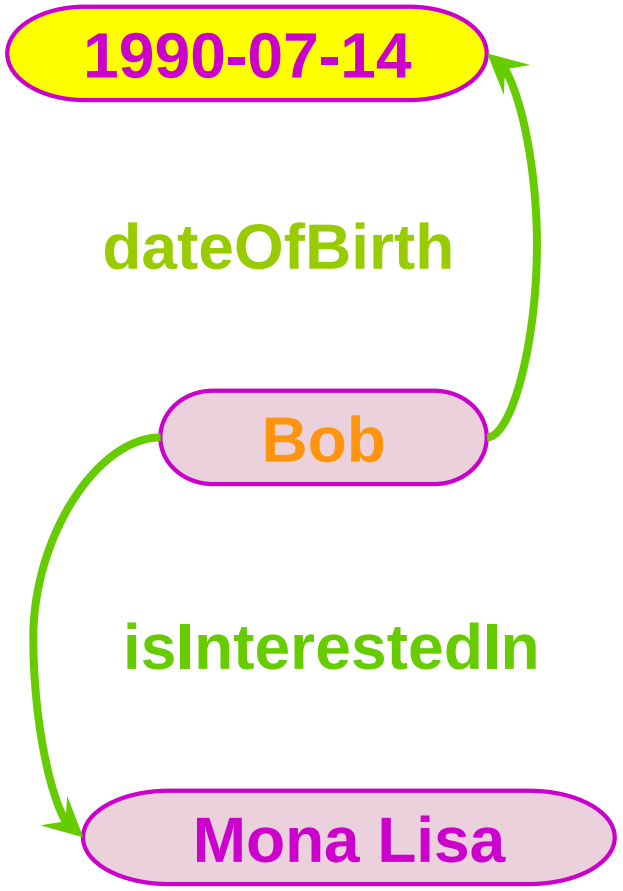
LABORATOIRE HUBERT CURIEN
UMR • CNRS • 5516 • SAINT-ETIENNE

An RDF graph


Alice
is a friend of
BOB
is interested in
is a
Person
is born on
14 July 1990
Leonardo Da Vinci
was created by
The Mona Lisa
is about
La Joconde à Washington

An RDF graph = A set of triples

(Subject , Predicate , Object)

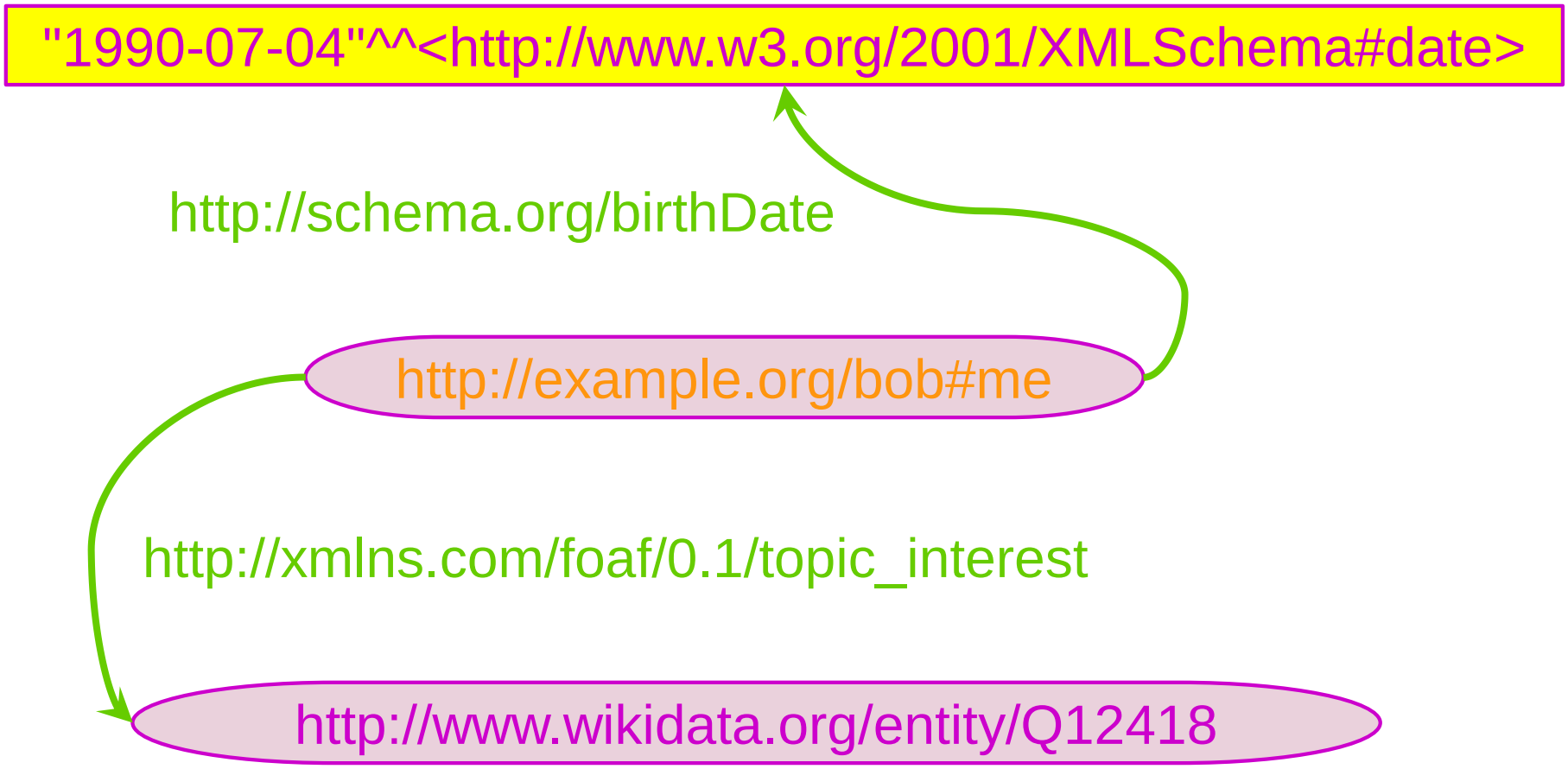


+URIs

nodes and arcs
uniquely identified

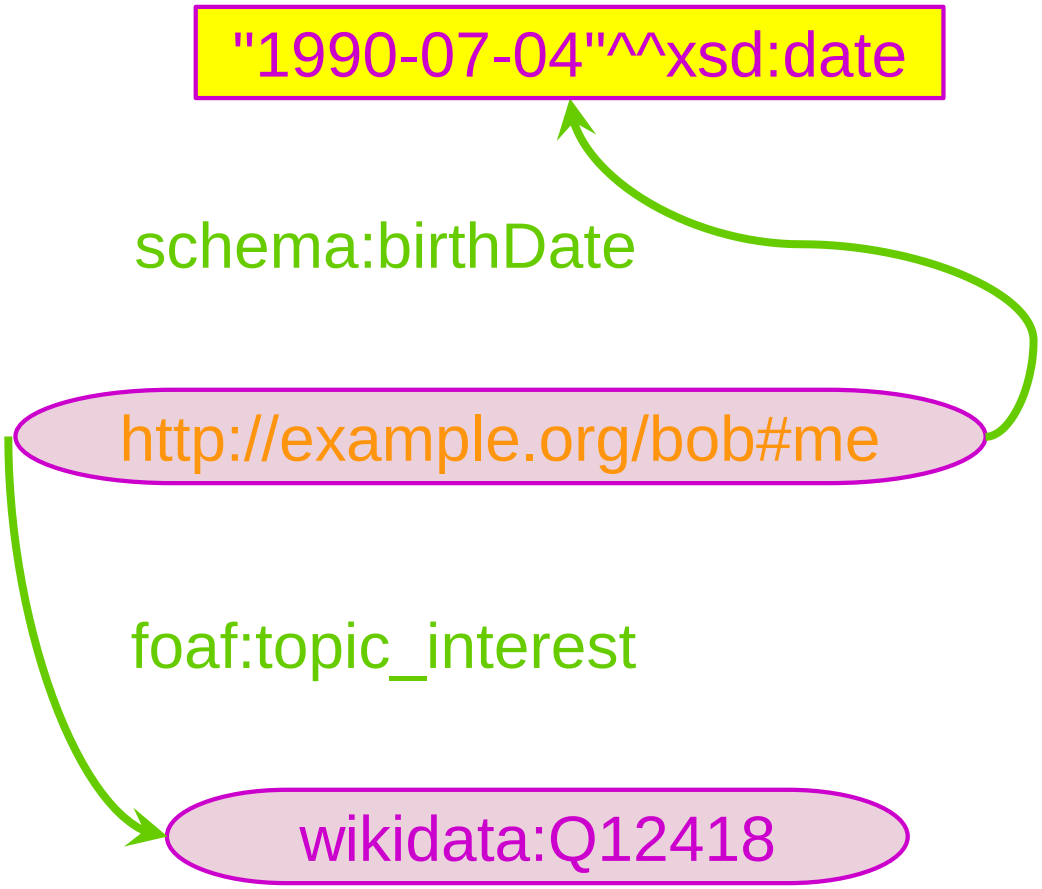
Resources: URIs, Literals, (+Blank nodes)

Subject:	URI or Blank node
Predicate:	URI
Object:	URI or Blank node or Literal

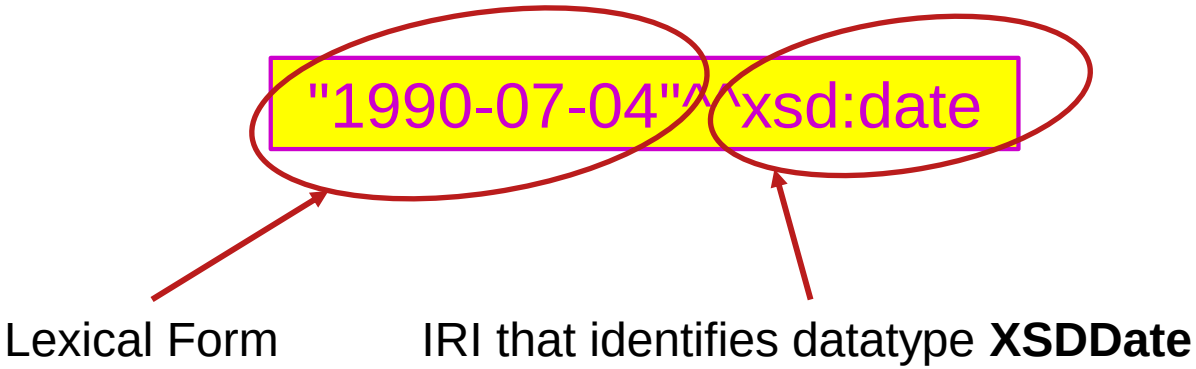


URI Prefixes

```
@Prefix xsd:       <http://www.w3.org/2001/XMLSchema#>.
@Prefix schema:    <http://schema.org/>.
@Prefix foaf:      <http://xmlns.com/foaf/0.1/>.
@Prefix wikidata: <http://www.wikidata.org/entity/>.
```



Literals and Datatypes



A Datatype *D*:

- a set *L(D)* of Unicode strings called the *Lexical Space*;
- a set *V(D)* called the *Value Space;*
- a mapping *L2V(D) : L(V)* → *V(D),* called the Lexical-to-value mapping, maps all Strings in *L(D)* to a value in *V(D)*

Heavily used datatypes

XSD datatypes **(W3C XSD 1.1 Part 2: Datatypes)**

3 Built-in Datatypes and Their Definitions
- 3.1 Namespace considerations
- 3.2 Special Built-in Datatypes
  - anySimpleType · anyAtomicType
- 3.3 Primitive Datatypes
  - string · boolean · decimal · float · double · duration · dateTime · time · date · gYearMonth · gYear · gMonthDay · gDay · gMonth · hexBinary · base64Binary · anyURI · QName · NOTATION
- 3.4 Other Built-in Datatypes
  - normalizedString · token · language · NMTOKEN · NMTOKENS · Name · NCName · ID · IDREF · IDREFS · ENTITY · ENTITIES · integer · nonPositiveInteger · negativeInteger · long · int · short · byte · nonNegativeInteger · unsignedLong · unsignedInt · unsignedShort · unsignedByte · positiveInteger · yearMonthDuration · dayTimeDuration · dateTimeStamp

GeoSPARQL wktLiterals **(OGC)**

```
"POLYGON((-77.050125 38.892086, -77.039482 38.892036, -77.039482 38.895393,
-77.033669 38.895508, -77.033585 38.892052, -77.031906 38.892086, -77.031883
38.887474, -77.050232 38.887142, -77.050125 38.892086 ))"^^geo:wktLiteral.
```

Quantity and Quantity values

CIM NGSI-LD INFORMATION MODEL

```
[
{"@id": "ParkingA",
"isadjacentTo": {"hasObject": {"@id":"StreetA"},
"hasOpening": {"@id":"GateA"}},
"parkingLastMaximumOcuppancy": "25 spaces"
"hasDirectPart": {"@id":"ParkingSpaceAH1"},
"hasDirectPart": {"@id":"ElectricChargingSpaceAE1"}},

{"@id": "StreetA",
"hasState": {"hasValue": "30% congested",
"reliability": "90%"},
"isConnectedTo": {"hasObject": {"@id":"AlleyB"},
"inDirection": " one way"}},

{"@id": "ParkingSpaceAH1",
"availablity": "available",
"applicableForUser": "handicapped"},

{"@id": "ElectricChargingSpaceAE1",
"availablity": "available",
"maximumCurrent": "50 Amperes"},
```

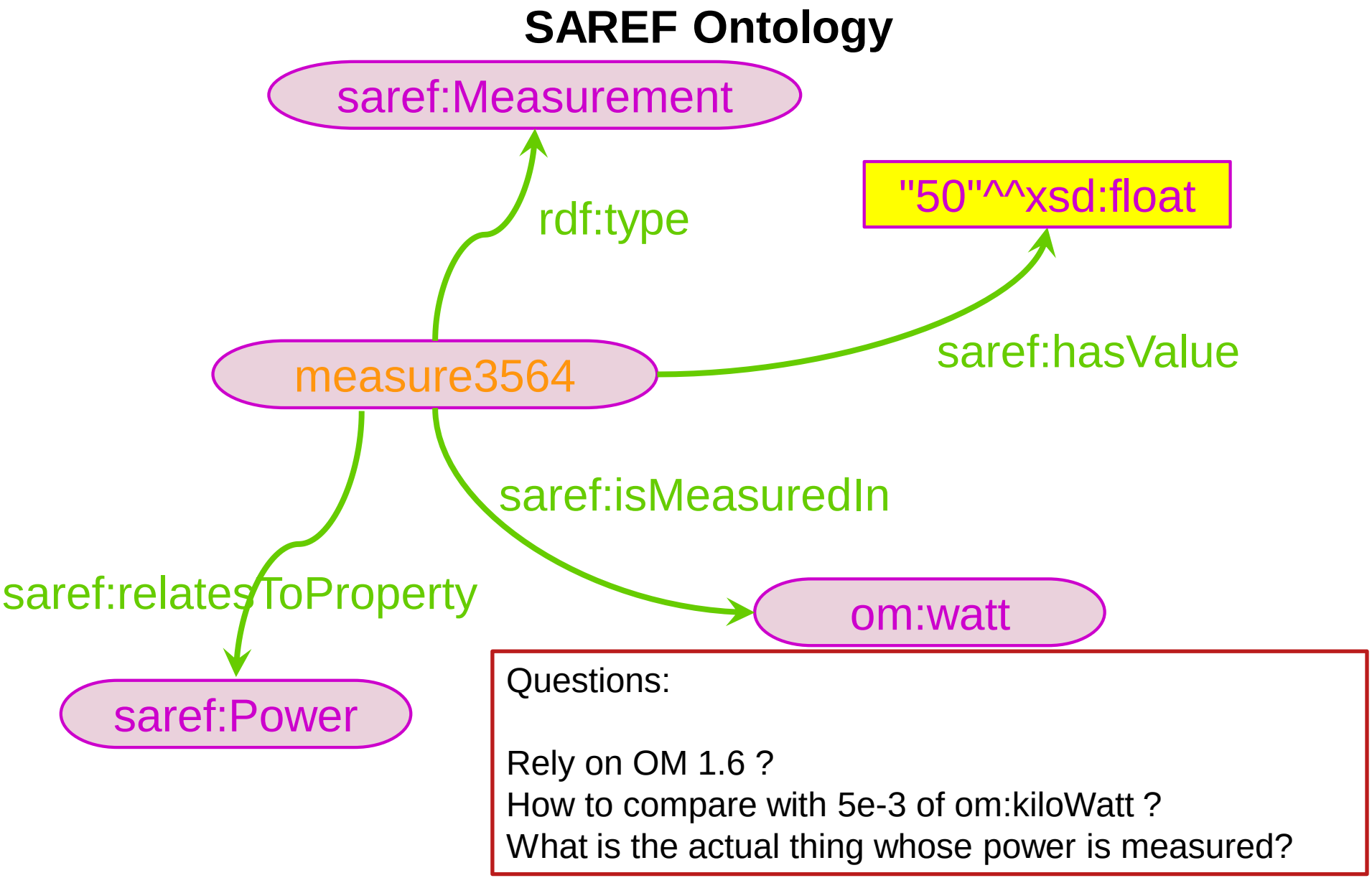
Questions:

What is the semantics of 50 Amperes?

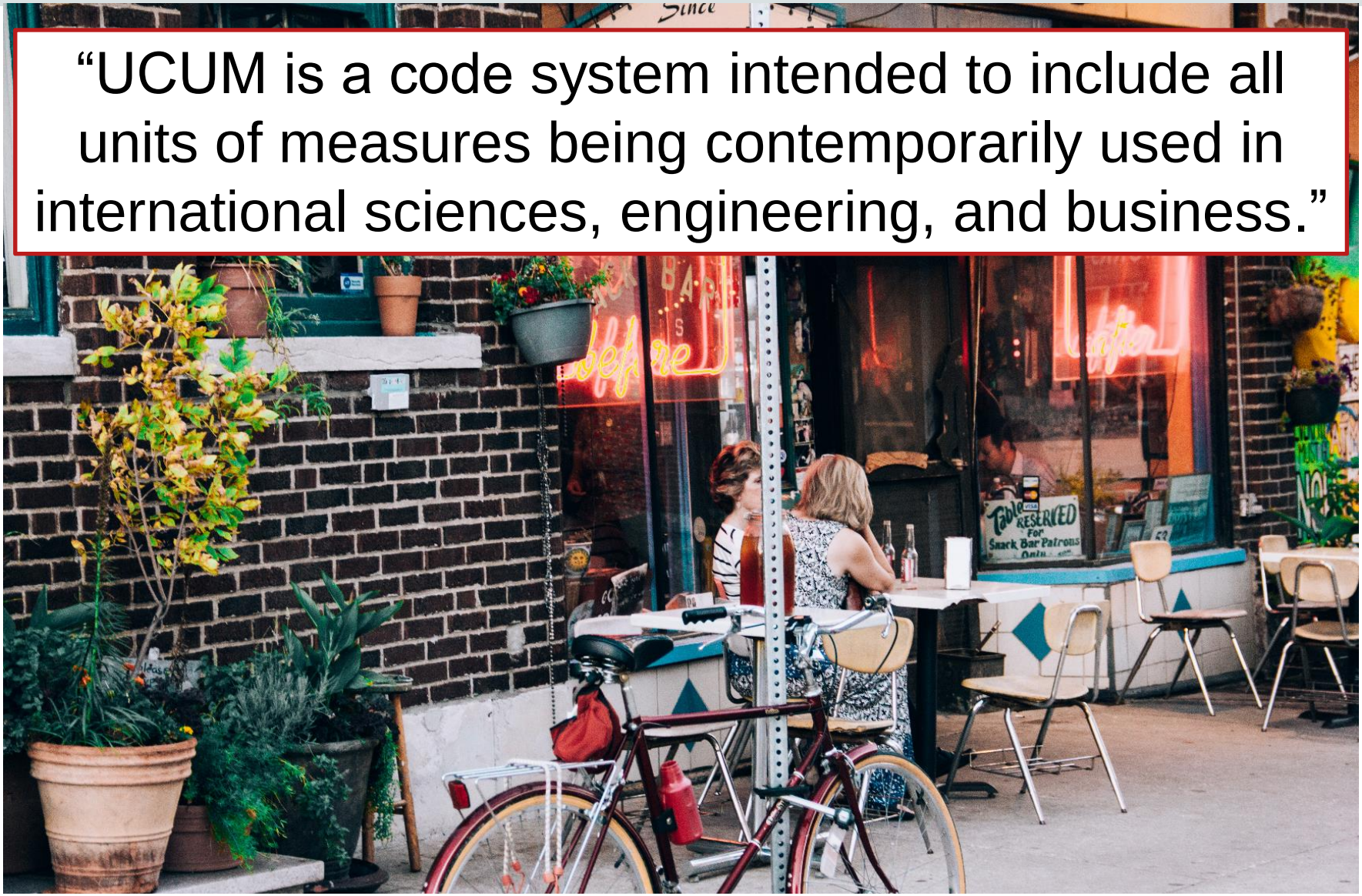
Are these values the same?

"50 A" "50 Ampere" "50 ampères" "50 Amperes" "50 000 mA"

Quantity and Quantity values



The cdt:ucum datatype

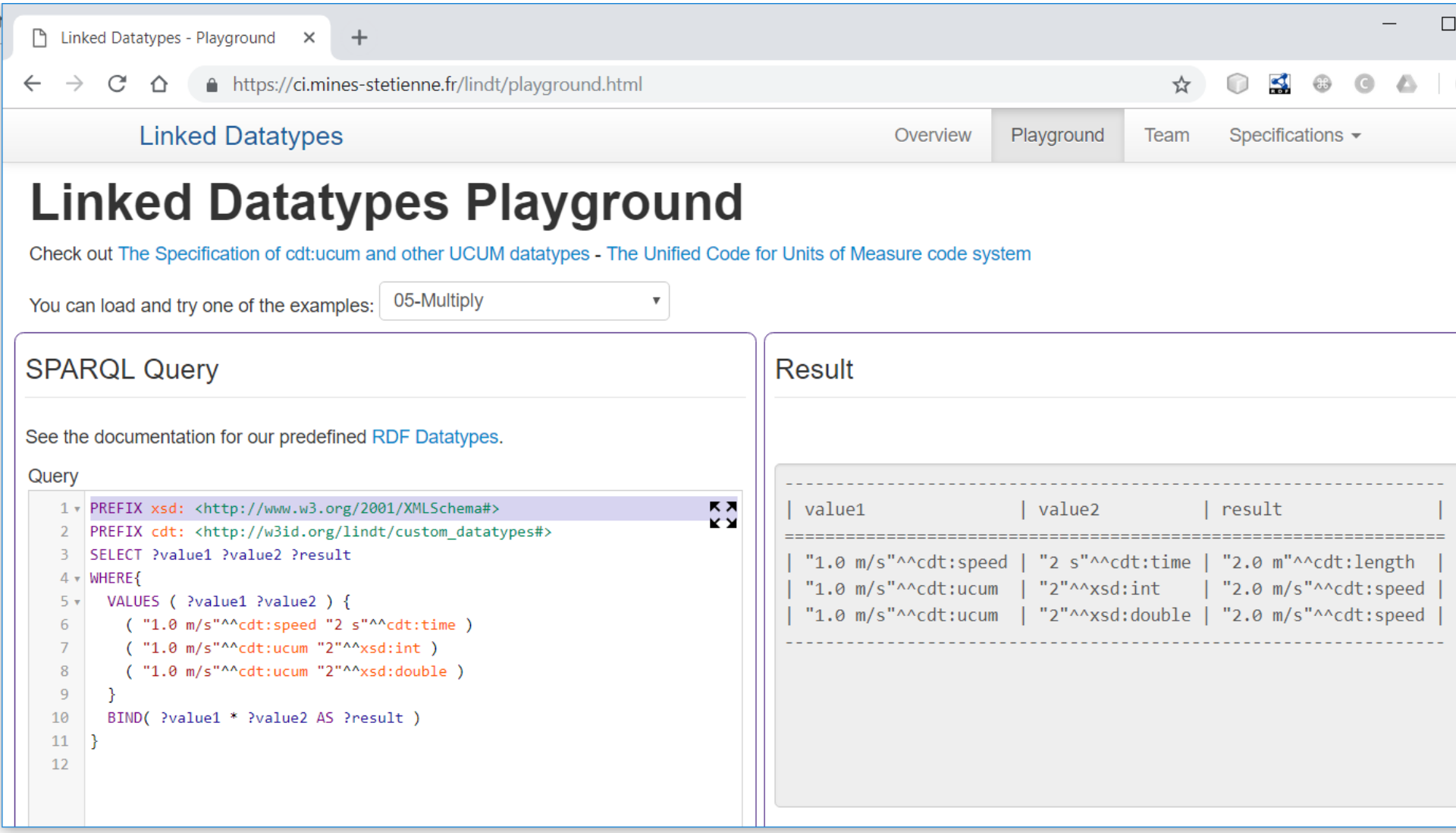



```
@prefix cdt: <http://w3id.org/lindt/custom_datatypes#>.
```

The cdt:ucum datatype

@prefix cdt: <http://w3id.org/lindt/custom_datatypes#>.

The cdt:ucum datatype



Conclusion

Quantity values are very important to represent context information, measurement, device metadata.

Quantity values are best modeled in RDF using well typed literals

The cdt:ucum datatype is designed to model quantity values

The cdt:ucum datatype uses UCUM (infinite number of units)

UCUM and cdt:ucum have open-source implementations