Institut Mines-Télécom

Ontologies, langages, tools, to ease the usage of RDF as a lingua franca

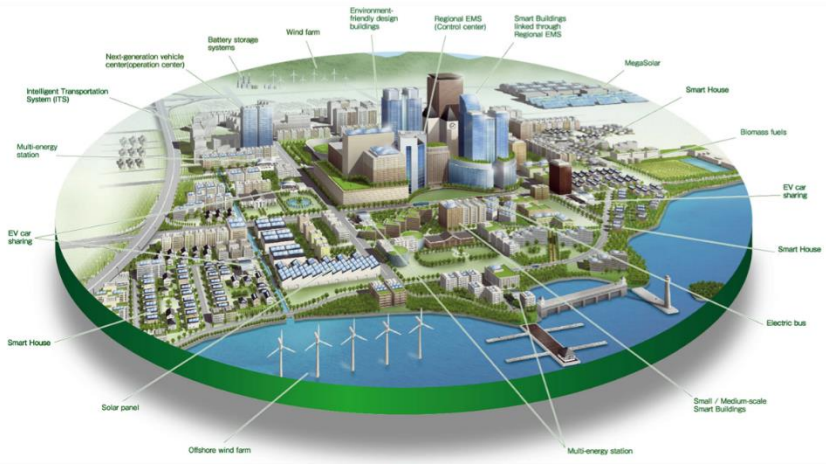
Maxime Lefrançois
Maxime.Lefrancois@emse.fr
http://maxime-lefrancois.info/

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

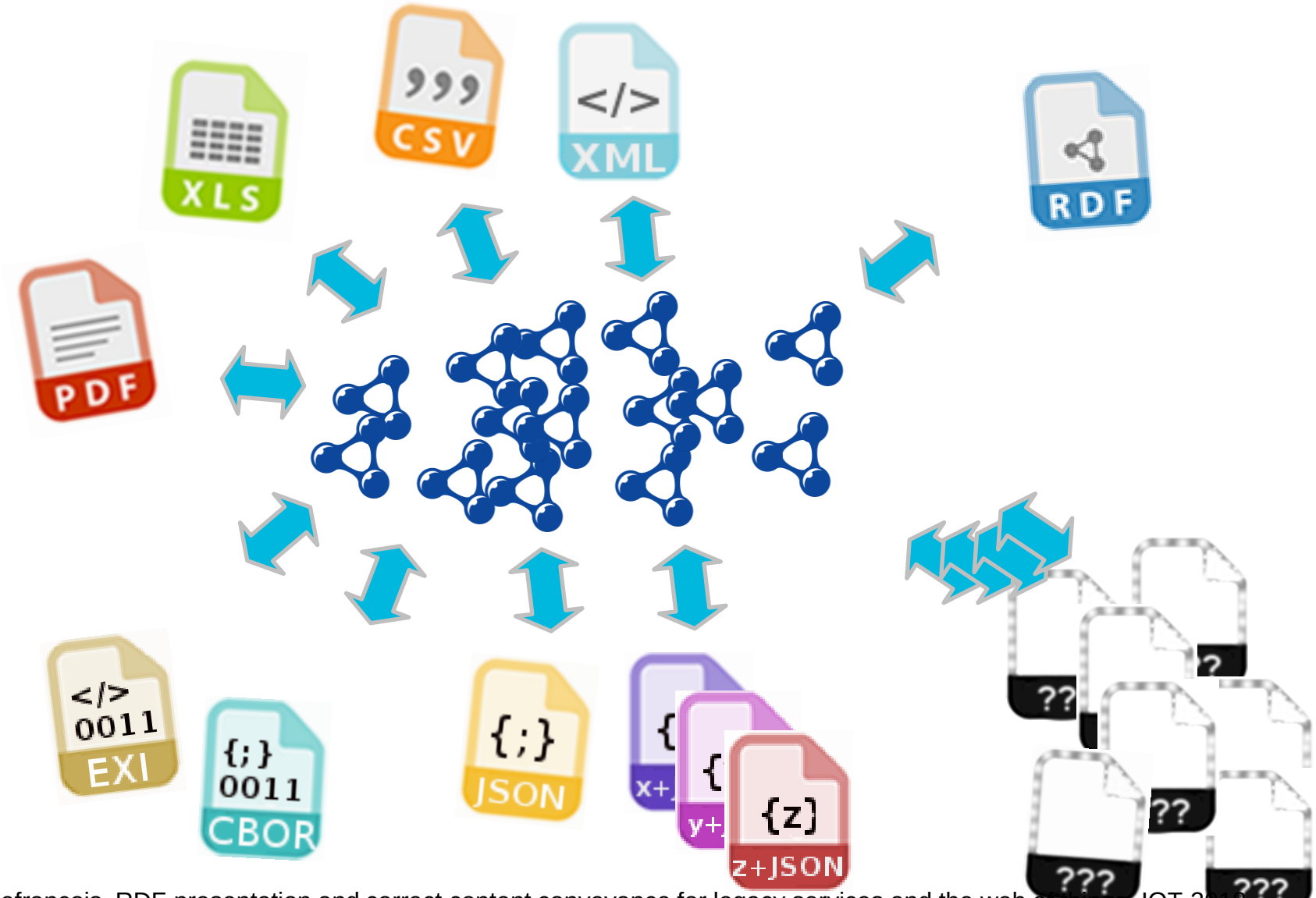
MINES Saint-Étienne
Une école de l'IMT

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

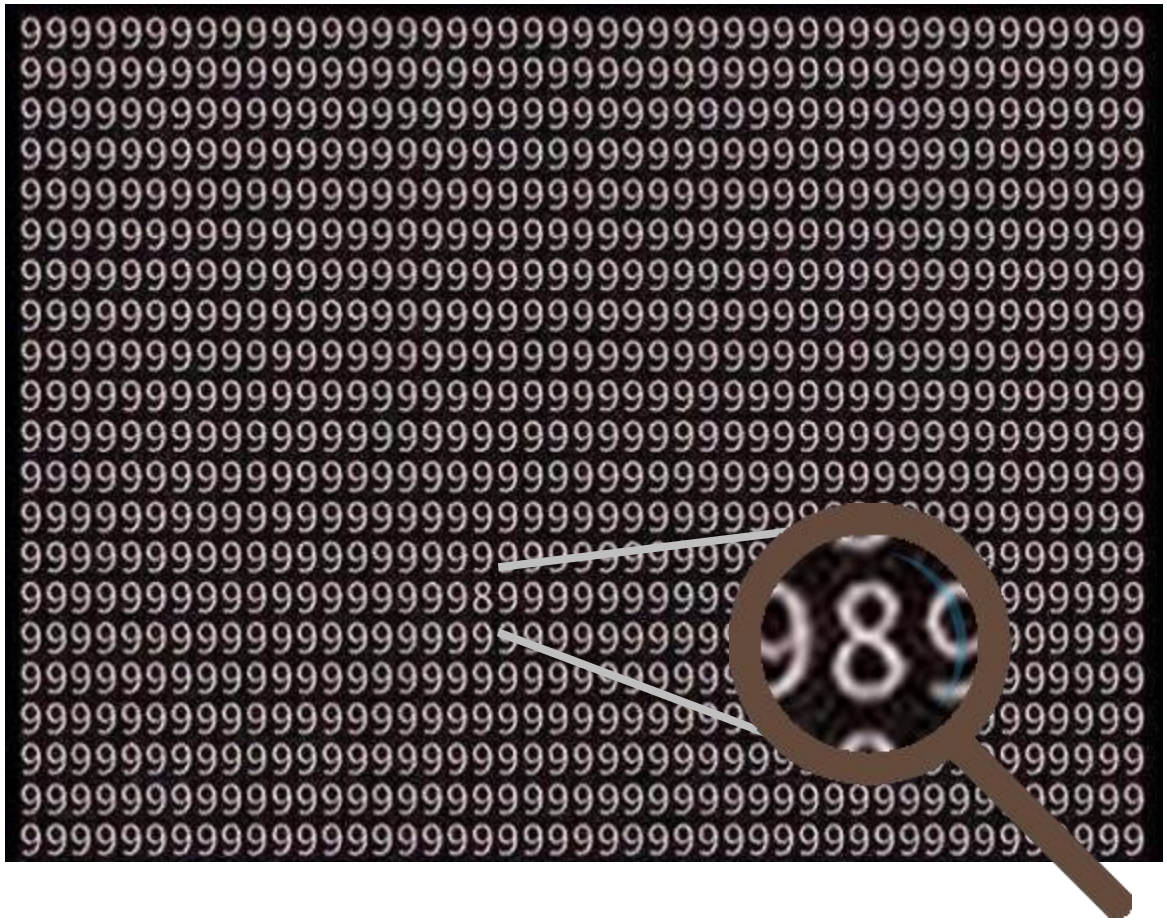
There is a multitude of data formats/models on the Web



Can we use RDF as a lingua franca?



Idea: just send RDF!



application/rdf+xml
text/turtle
application/ld+json
…
ERI
HDTQ

RDF data *formats* will never be the only ones on the web
Developers prefer JSON, …

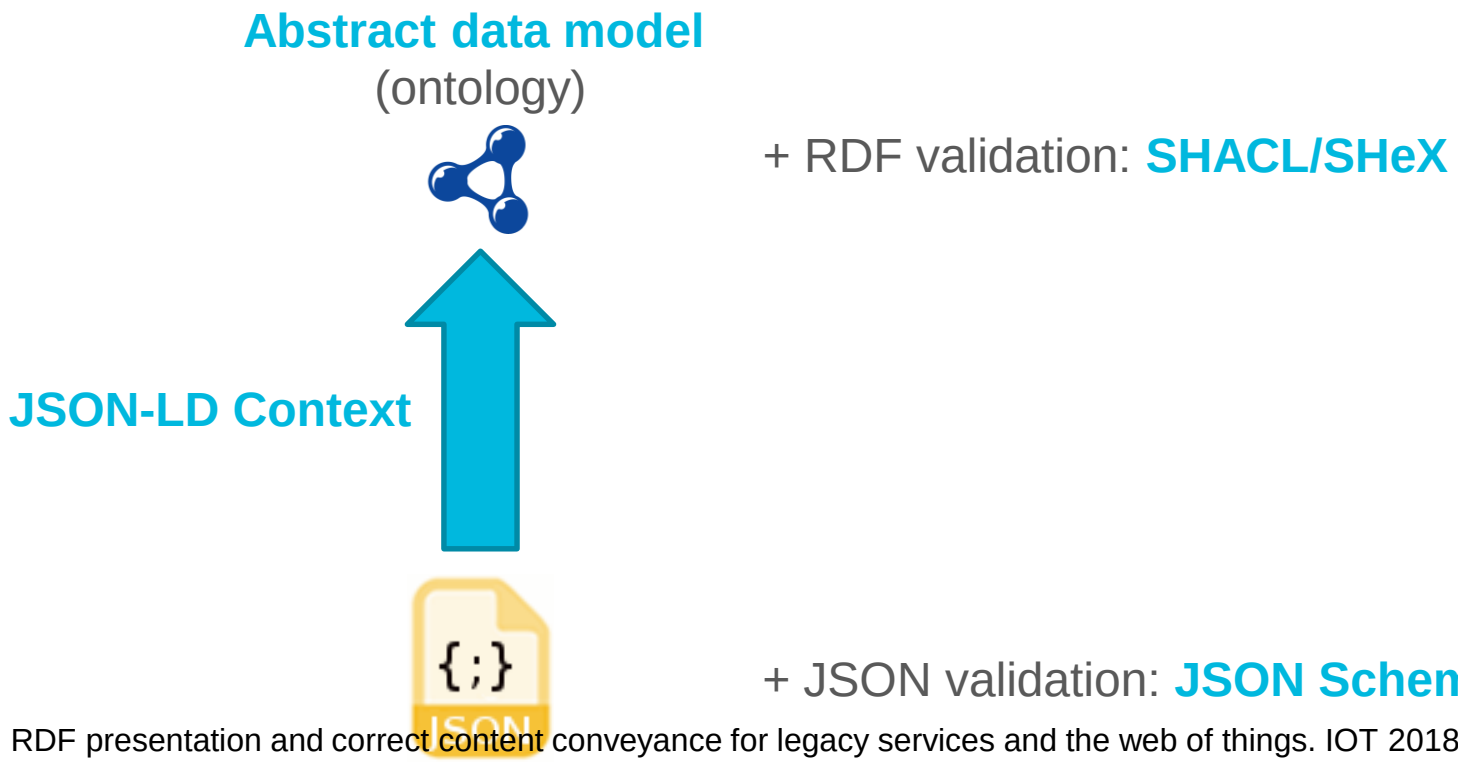
Idea: just send ~~RDF!~~ JSON-LD!

Solution:

- to rely on *one* RDF syntax like JSON-LD

Requires:

- Global adoption of JSON-LD (Utopian)
- Maintaining during the development or the evolution phase



Approx. modeling of the communication between heterogeneous agents on the Web.

sender

wants to send some content

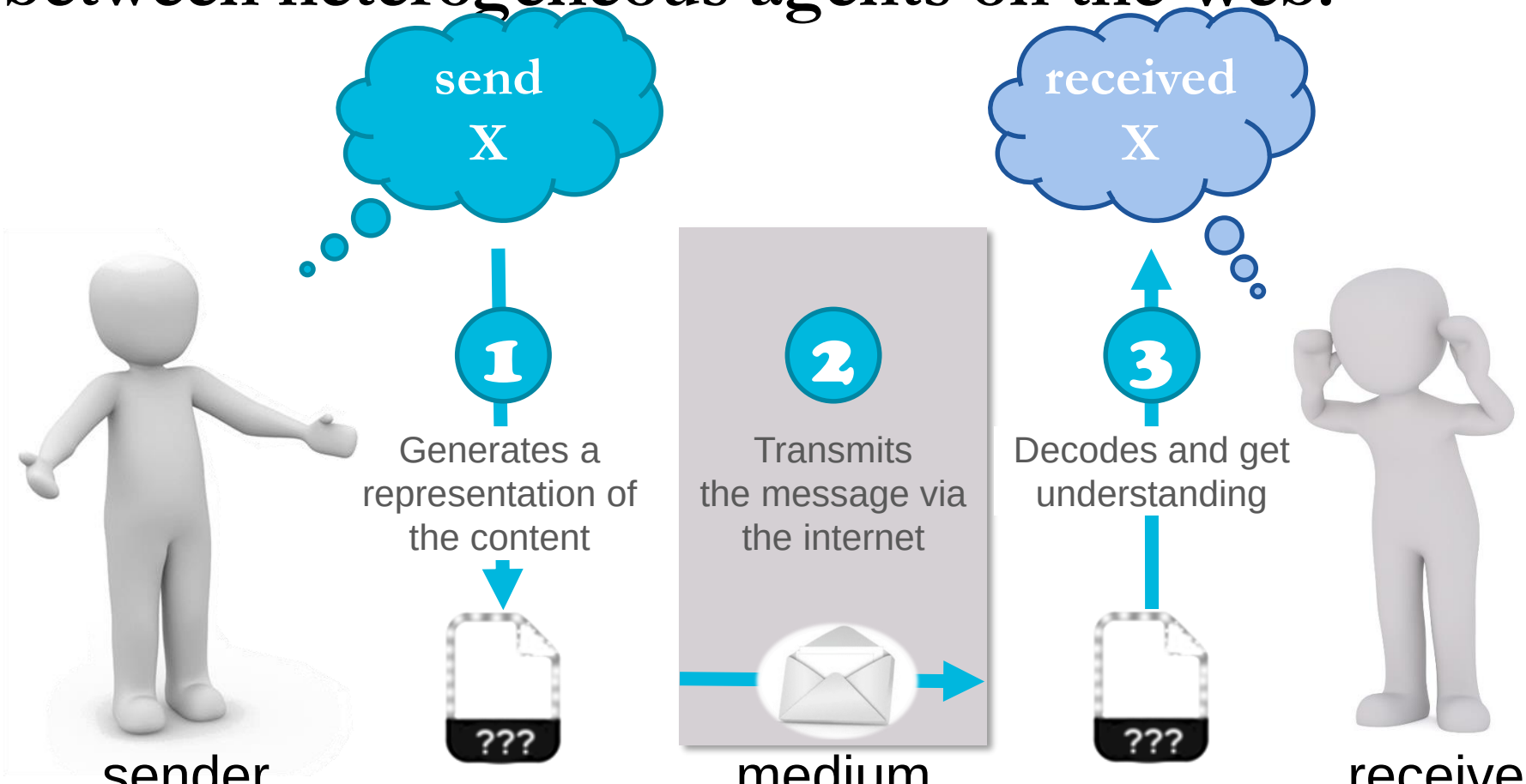
e.g., state of a resource

medium

receiver

M. Lefrançois, RDF presentation and correct content conveyance for legacy services and the web of things. IOT 2018

Approx. modeling of the communication between heterogeneous agents on the Web.

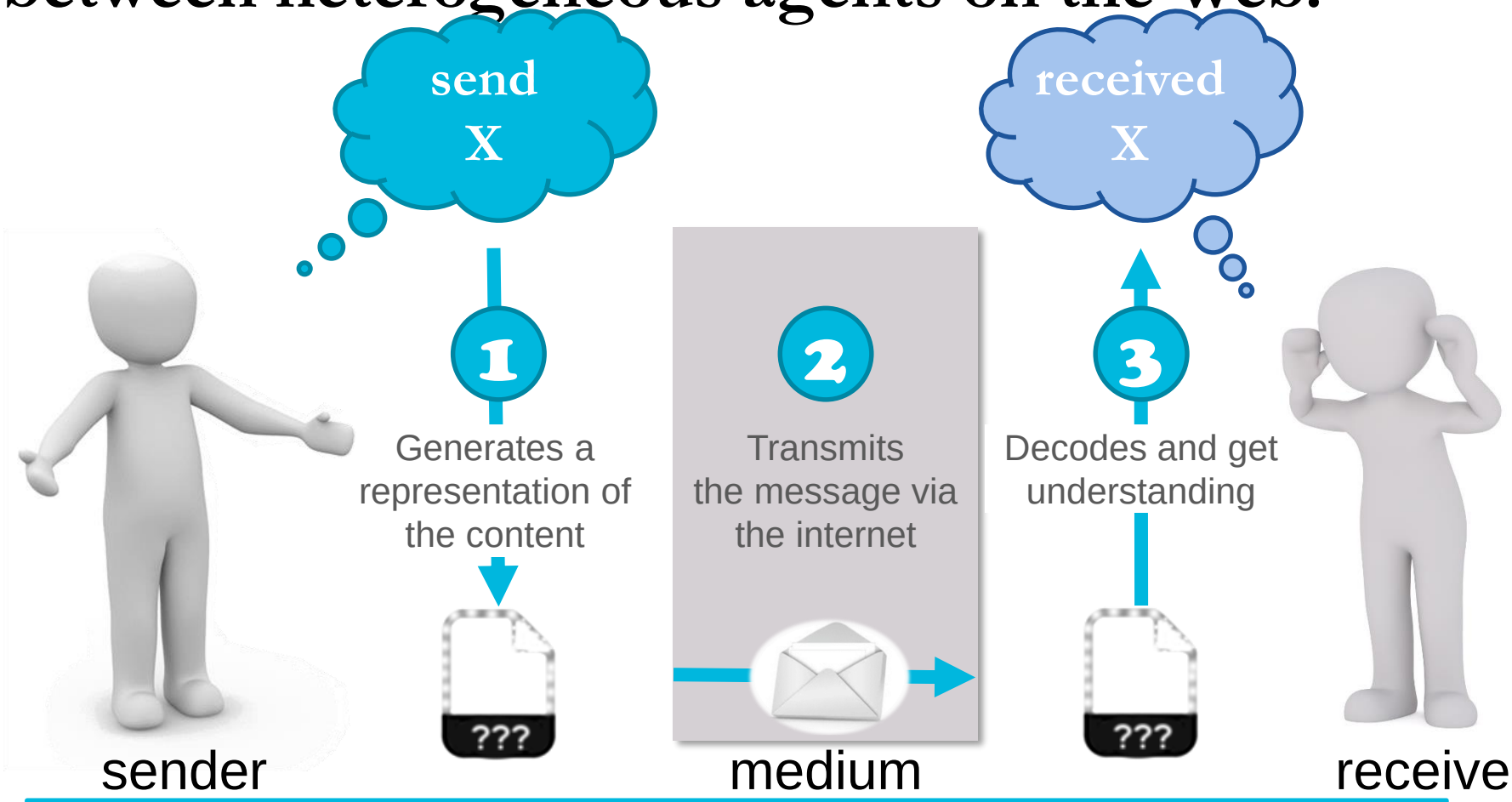


wants to send some content

e.g., state of a resource

M. Lefrançois, RDF presentation and correct content conveyance for legacy services and the web of things. IOT 2018

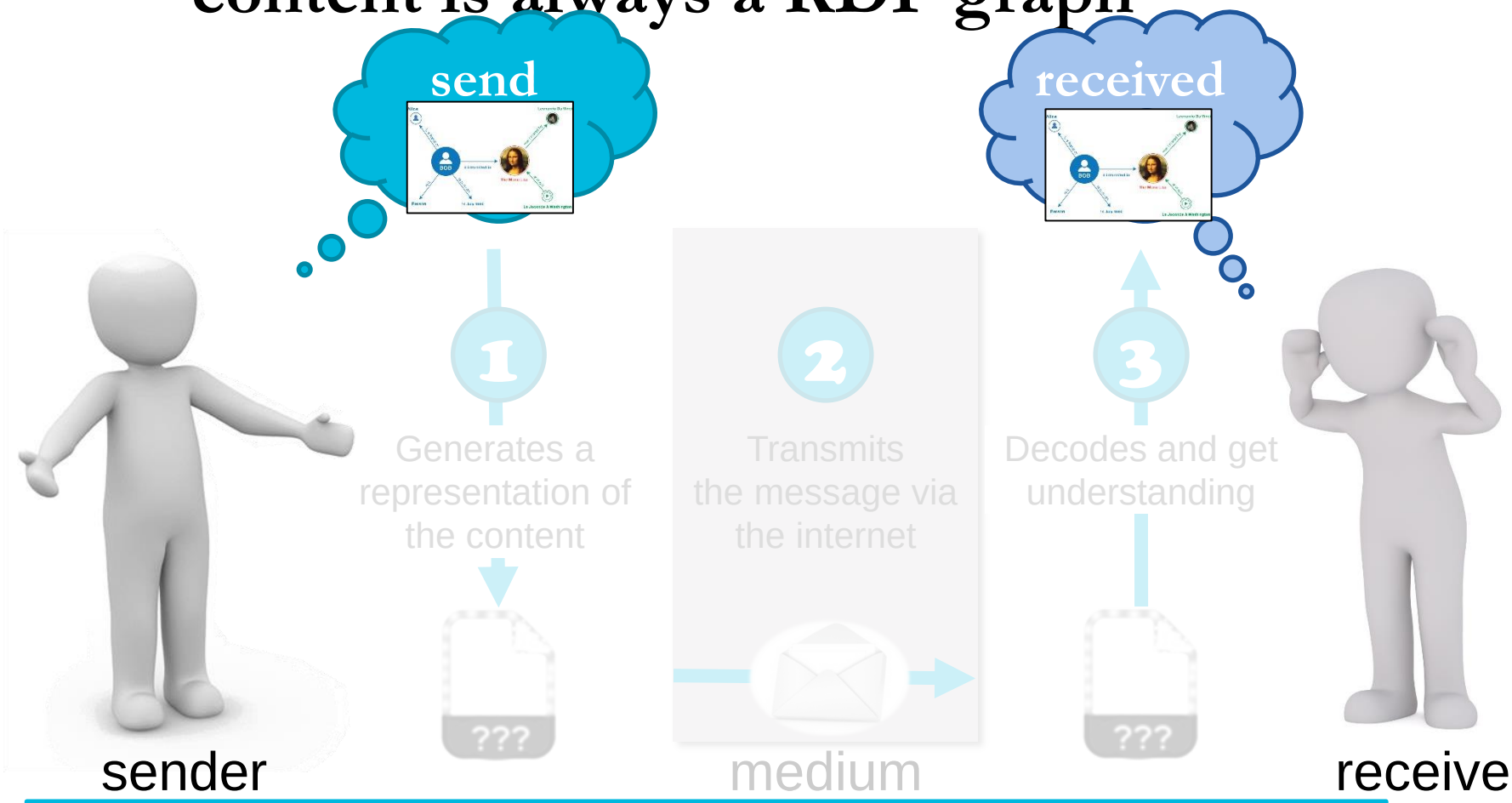
Approx. modeling of the communication between heterogeneous agents on the Web.



Correct Content Conveyance **iif:**

1. all of the essential characteristics of the content is encoded in the message
2. the encoding and the decoding phase are symmetric
3. the message is not altered in the transmission medium

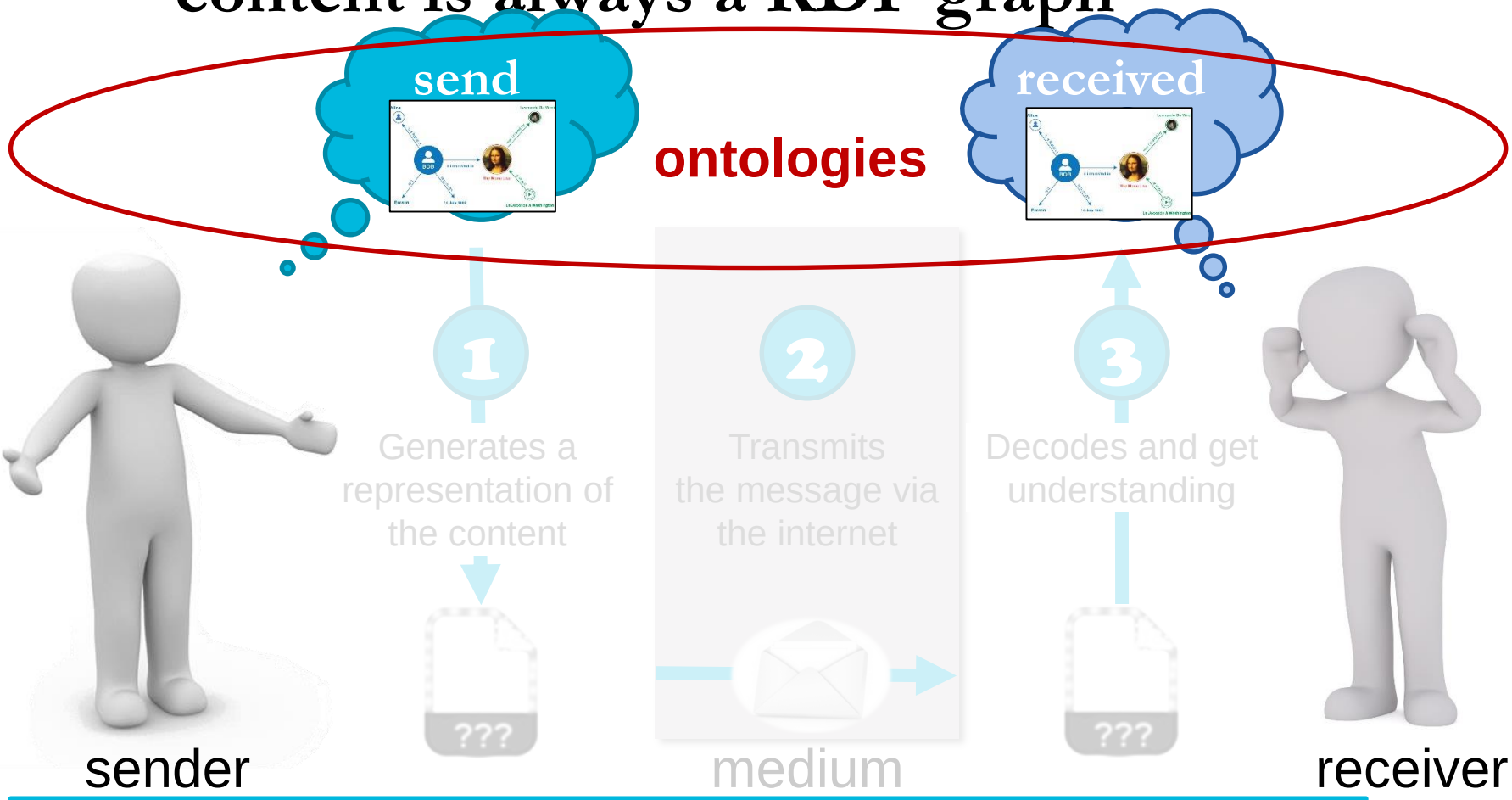
Assumption:
content is always a RDF graph



Correct Content Conveyance **iif:**

The RDF graph the sender encodes is equivalent to the RDF graph the receiver obtained after decoding the message

Assumption:
content is always a RDF graph



Correct Content Conveyance **iif:**

The RDF graph the sender encodes is equivalent to the RDF graph the receiver obtained after decoding the message

Ontologies, contribution to standardization
SOSA/SSN ontology

Standard OGC et W3C

W3C Recommendation

Semantic Sensor Network Ontology

OGC® Making location count. W3C

W3C Recommendation 19 October 2017 (Link errors corrected 08 December 2017)

This version:
https://www.w3.org/TR/2017/REC-vocab-ssn-20171019/

Latest published version:
https://www.w3.org/TR/vocab-ssn/

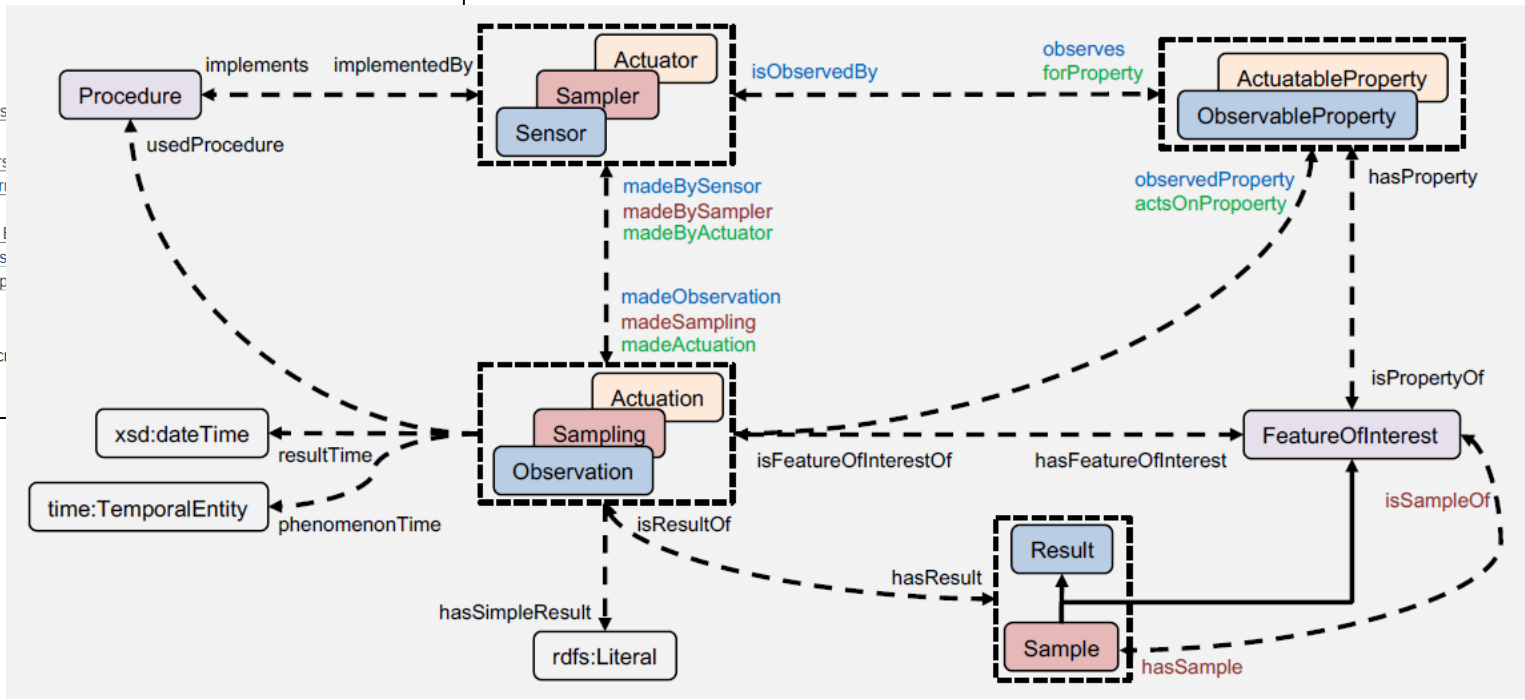
Latest editor's draft:
https://w3c.github.io/sdw/ssn/

Implementation report:
https://w3c.github.io/sdw/ssn-usage/

Previous version:
https://www.w3.org/TR/2017/PR-vocab-s

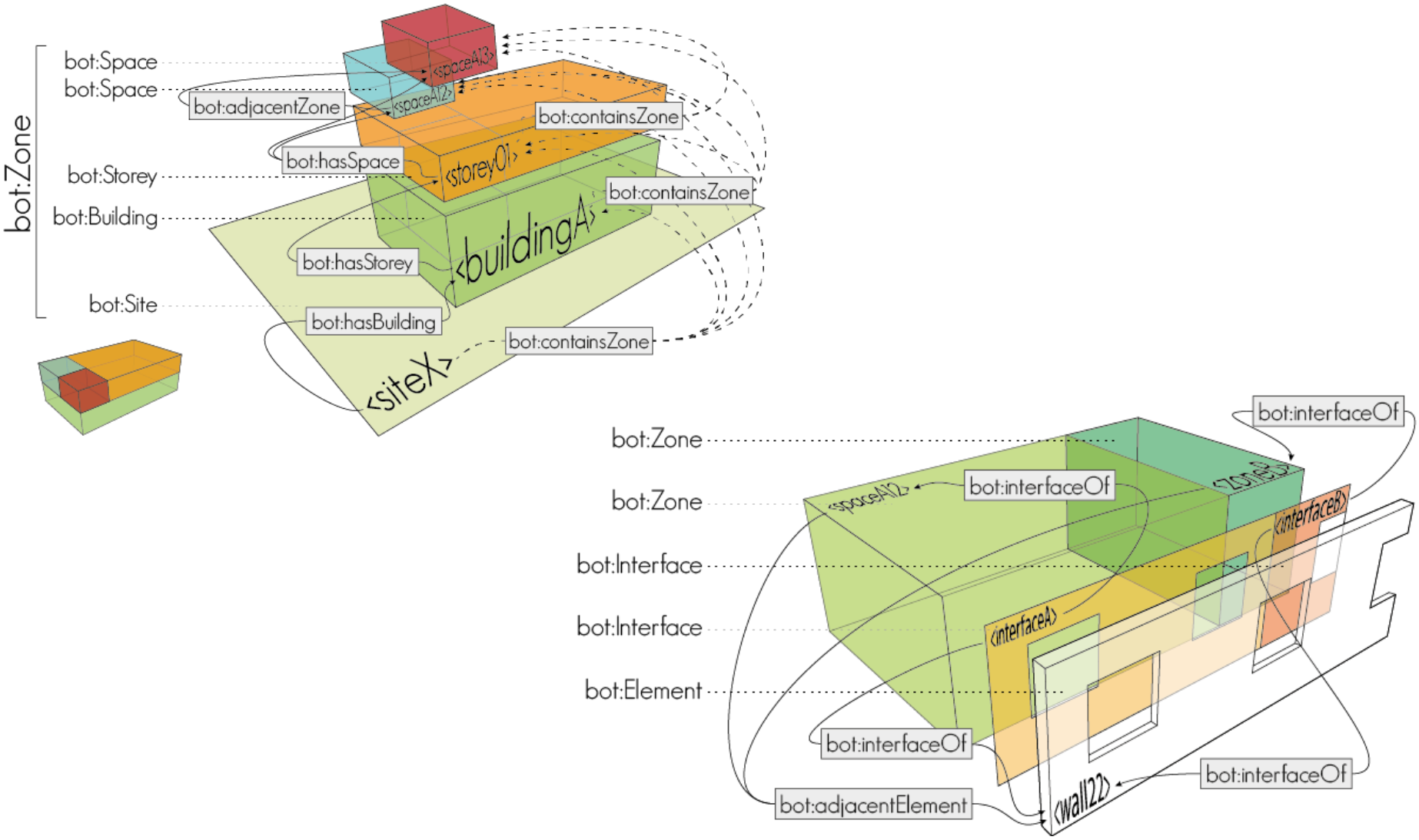
Editors:
Armin Haller, Australian National Univers
Krzysztof Janowicz, University of Califor
Simon Cox, CSIRO
Danh Le Phuoc, Technical University of
Kerry Taylor, Australian National Univers
Maxime Lefrançois, École Nationale Sup

Contributors (ordered alphabetically):
Rob Atkinson, Metalinkage
Raúl García-Castro, Universidad Politécn
Joshua Lieberman, Tumbling Walls
Claus Stadler, Universität Leipzig



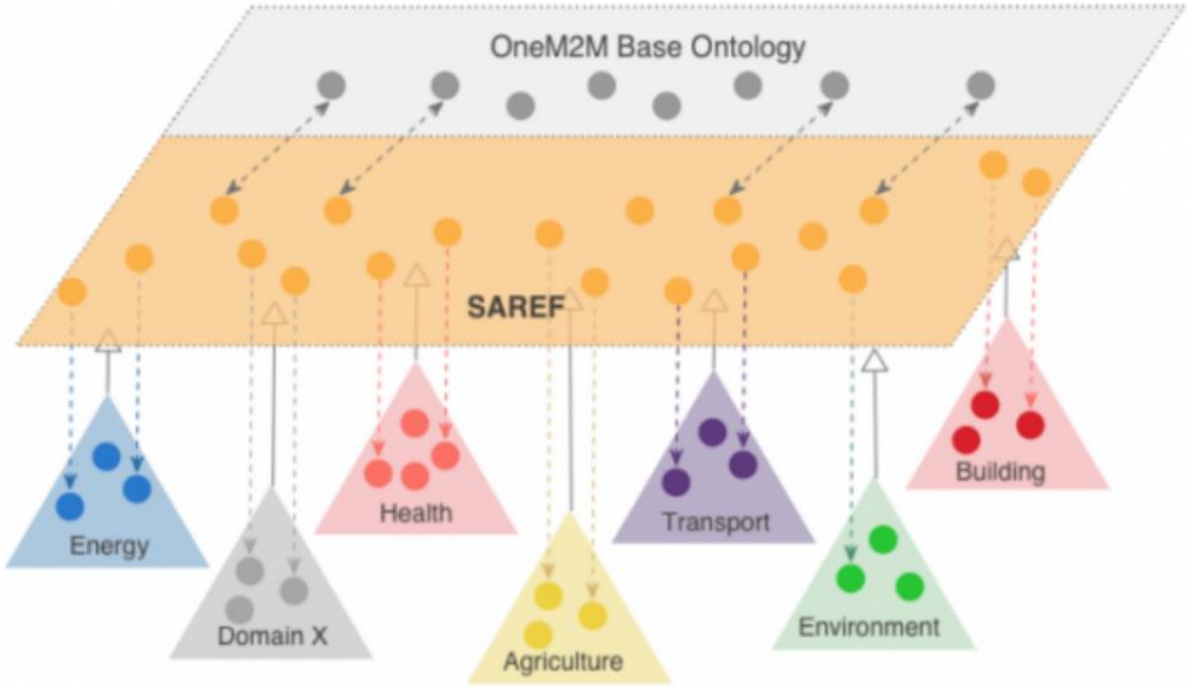
A. Haller, K. Janowicz, S. Cox, D. Le Phuoc, K. Taylor, and M. Lefrançois, Semantic Sensor Network Ontology, W3C Recommendation, W3C, 19 October 2017

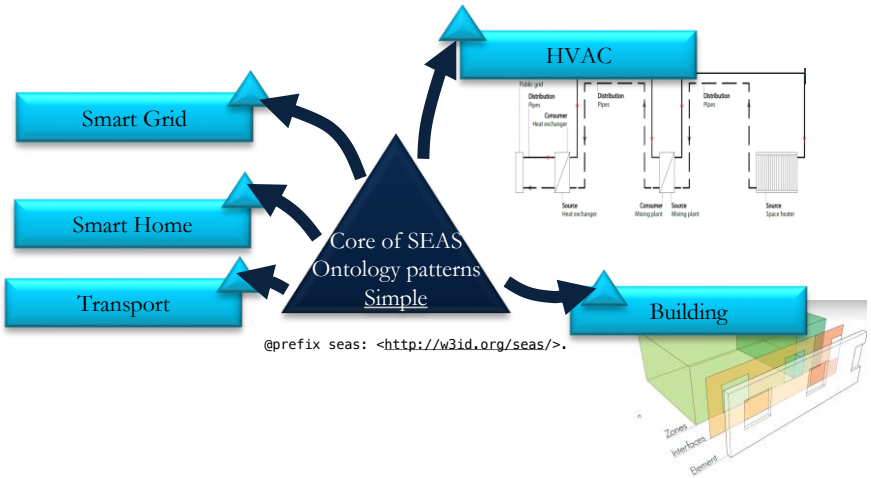
Ontologies, contribution to standardization
BOT ontology

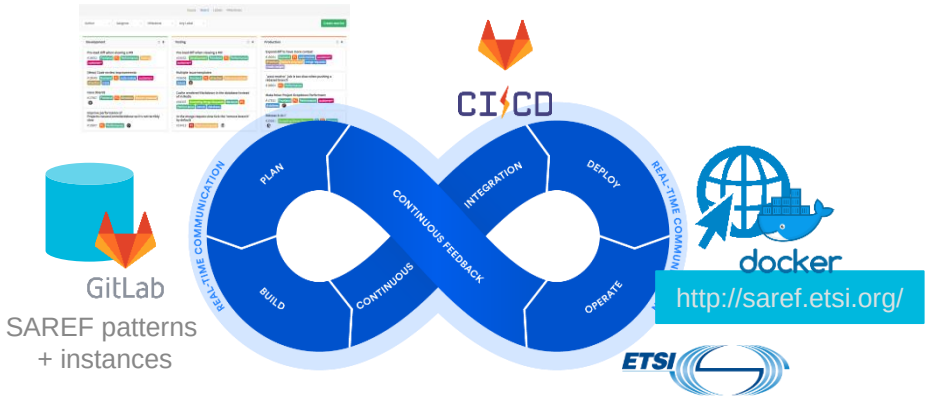


Mads Holten Rasmussen, Maxime Lefrançois, Georg Ferdinand Schneider, Pieter Pauwels, BOT: the Building Topology Ontology of the W3C Linked Building Data Group (submitted May 2019 to Semantic Web Journal)

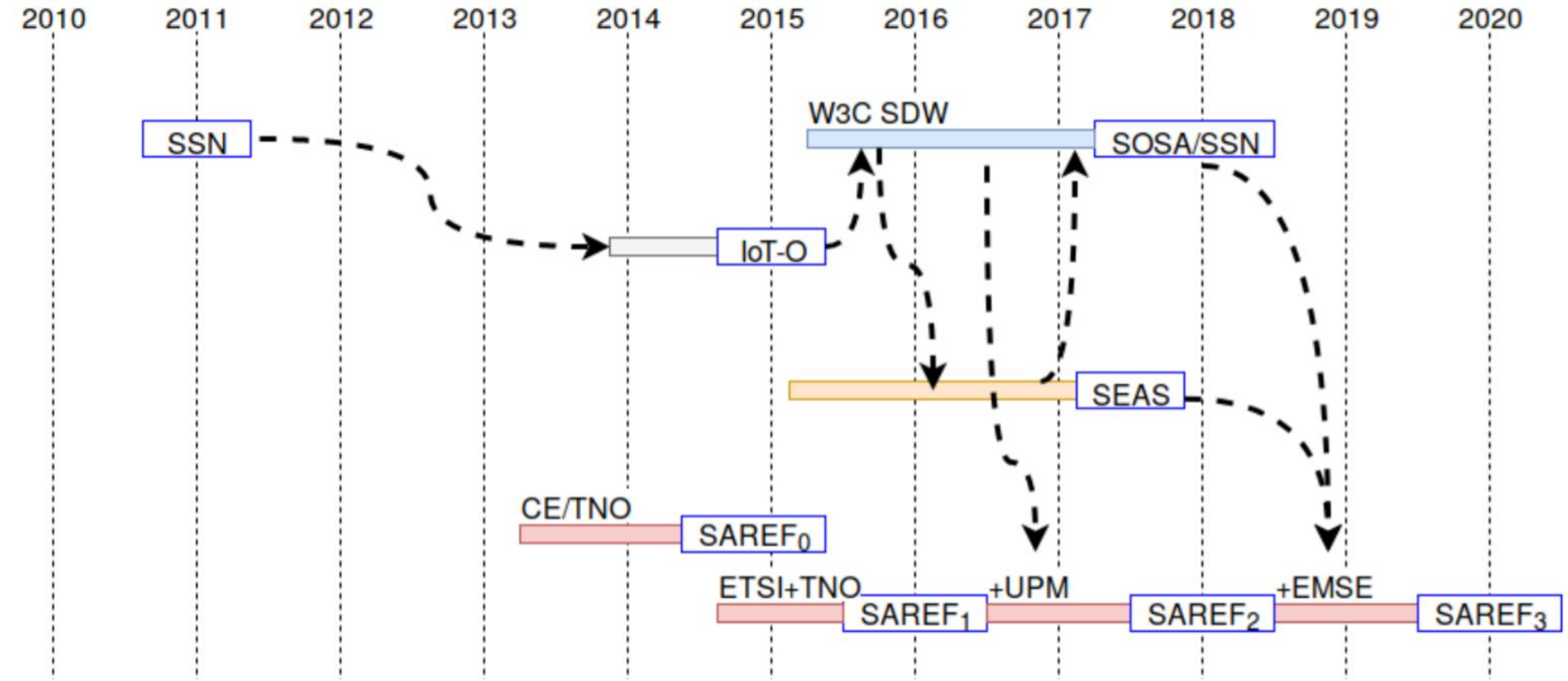
Ontologies, contribution to standardization
SAREF + influences



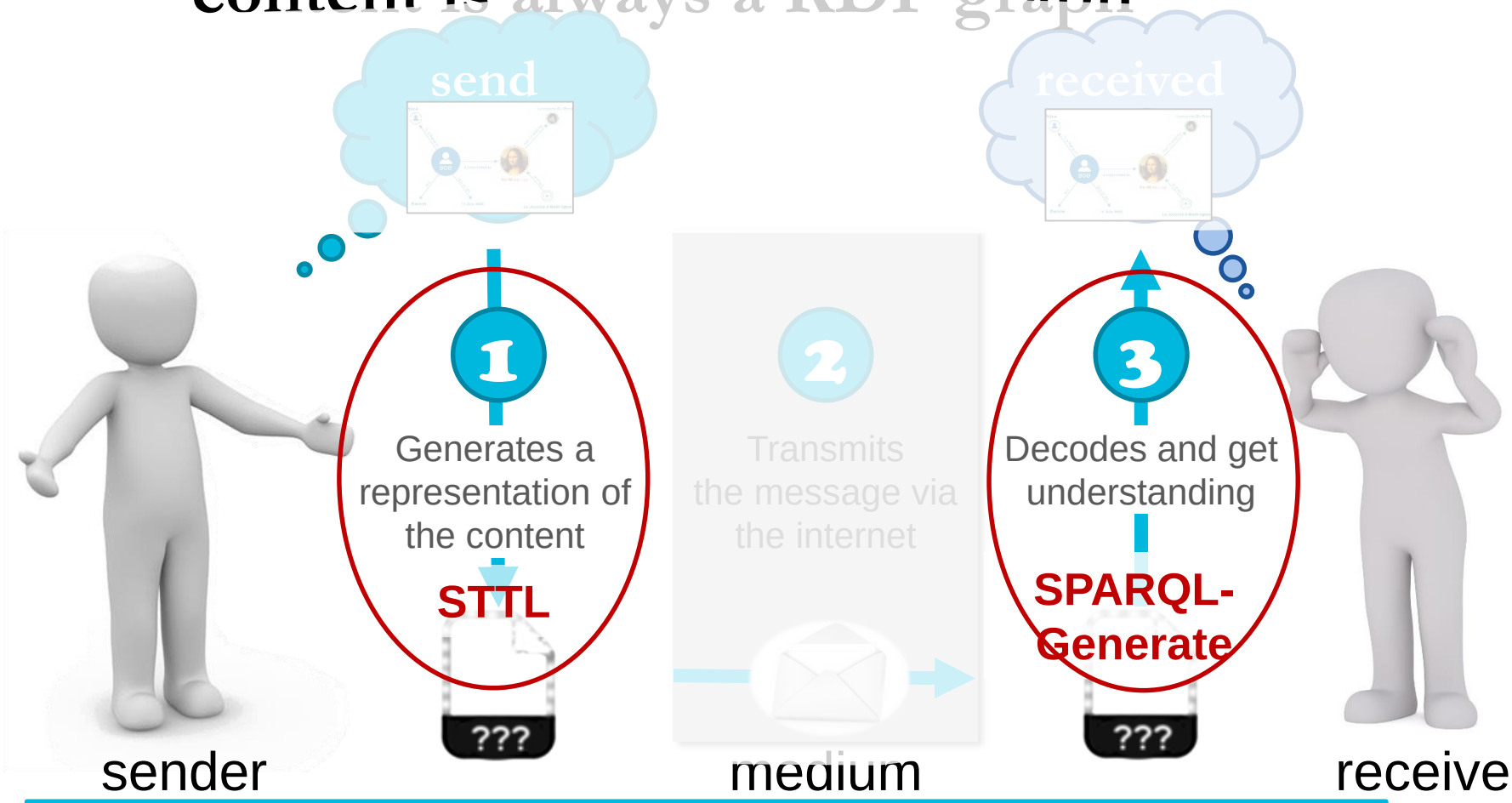




Ontologies, contribution to standardization
A work-in progress at international level



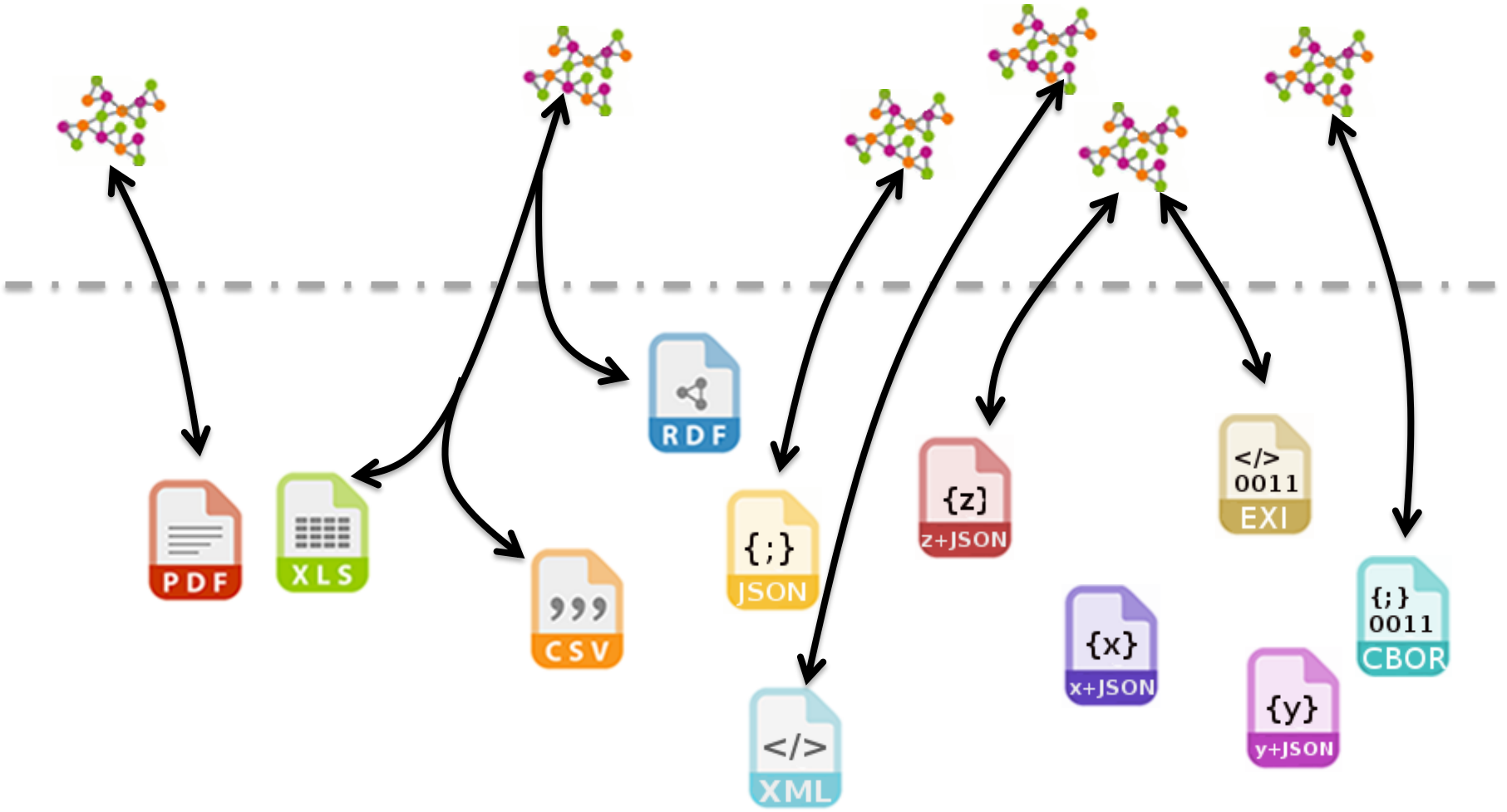
Assumption:
content is always a RDF graph



Correct Content Conveyance **iif:**

The RDF graph the sender encodes is equivalent to the RDF graph the receiver obtained after decoding the message

SPARQL-Generate
Heterogeneous data integration language



M. Lefrançois, ,A. Zimmermann, N. Bakerally, A SPARQL extension for generating RDF from heterogeneous formats, In Proc. Extended Semantic Web Conference, 2017
+ EKAW, Hackatons, Tutoriel à ESWC 2018

SPARQL-Generate
Heterogeneous data integration language

https://w3id.org/sparql-generate/

Query and generate both RDF and heterogeneous documents.

Generates RDF or text from documents
in XML, JSON, CSV, GeoJSON, HTML, CBOR, texte, …

Generates RDF or text streams from streams of
big CSV documents, MQTT streams, WebSocket, …

Open-source Licence Apache 2.0 implementation (financed by ITEA, ANR, ENGIE)

Expressive, Perfomant, extensible to other formats

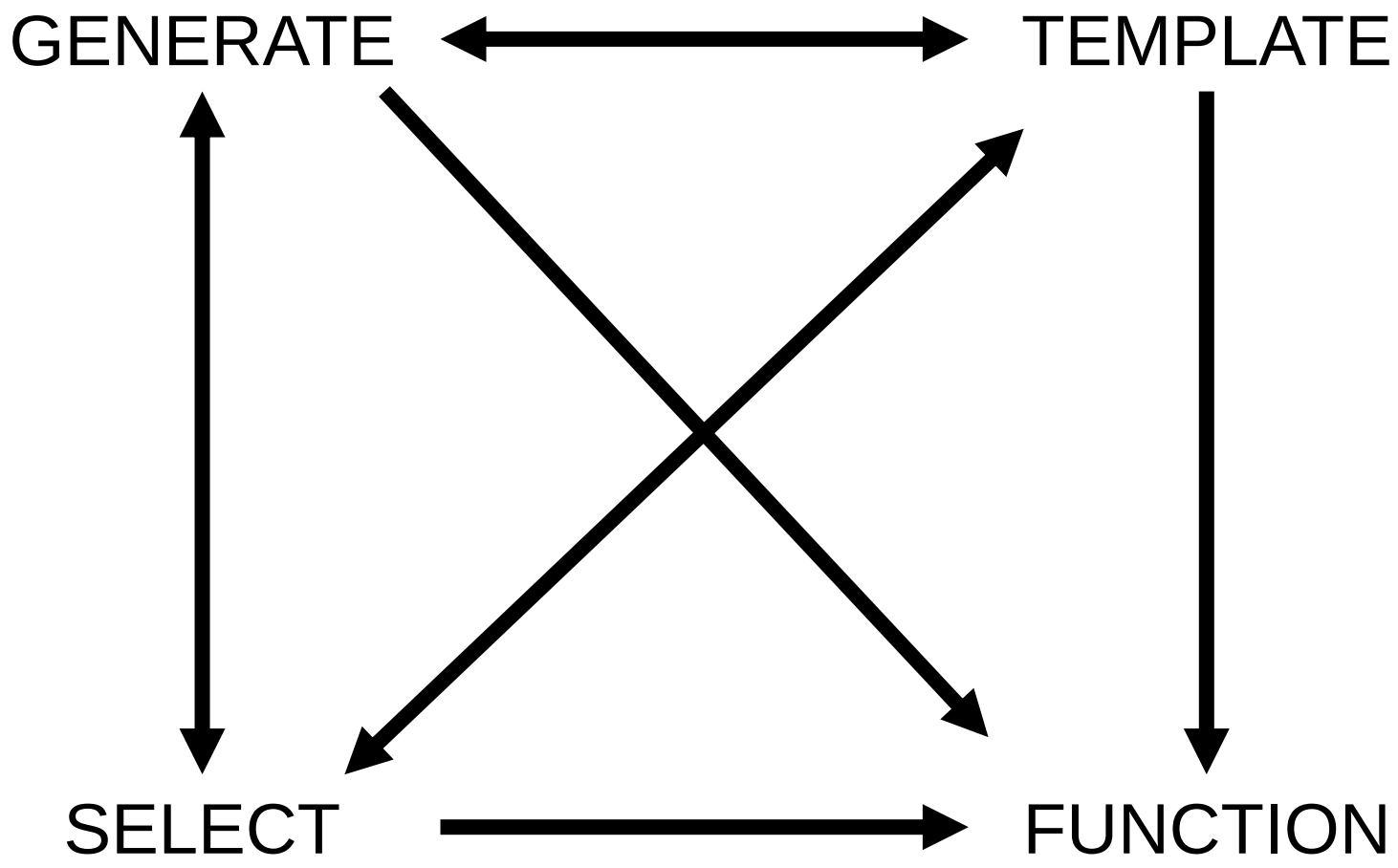
M. Lefrançois, ,A. Zimmermann, N. Bakerally, A SPARQL extension for generating RDF from heterogeneous formats, In Proc. Extended Semantic Web Conference, 2017
+ EKAW, Hackatons, Tutoriel à ESWC 2018

Additions in the past months/weeks

- Syntactic sugar
- Data streams
- More expressive fuctions
- RDF Lists
- Asynchronicity / multi-threading
- Aggregates
- SPARQL SELECT
- SPARQL FUNCTION
- SPARQL TEMPLATE

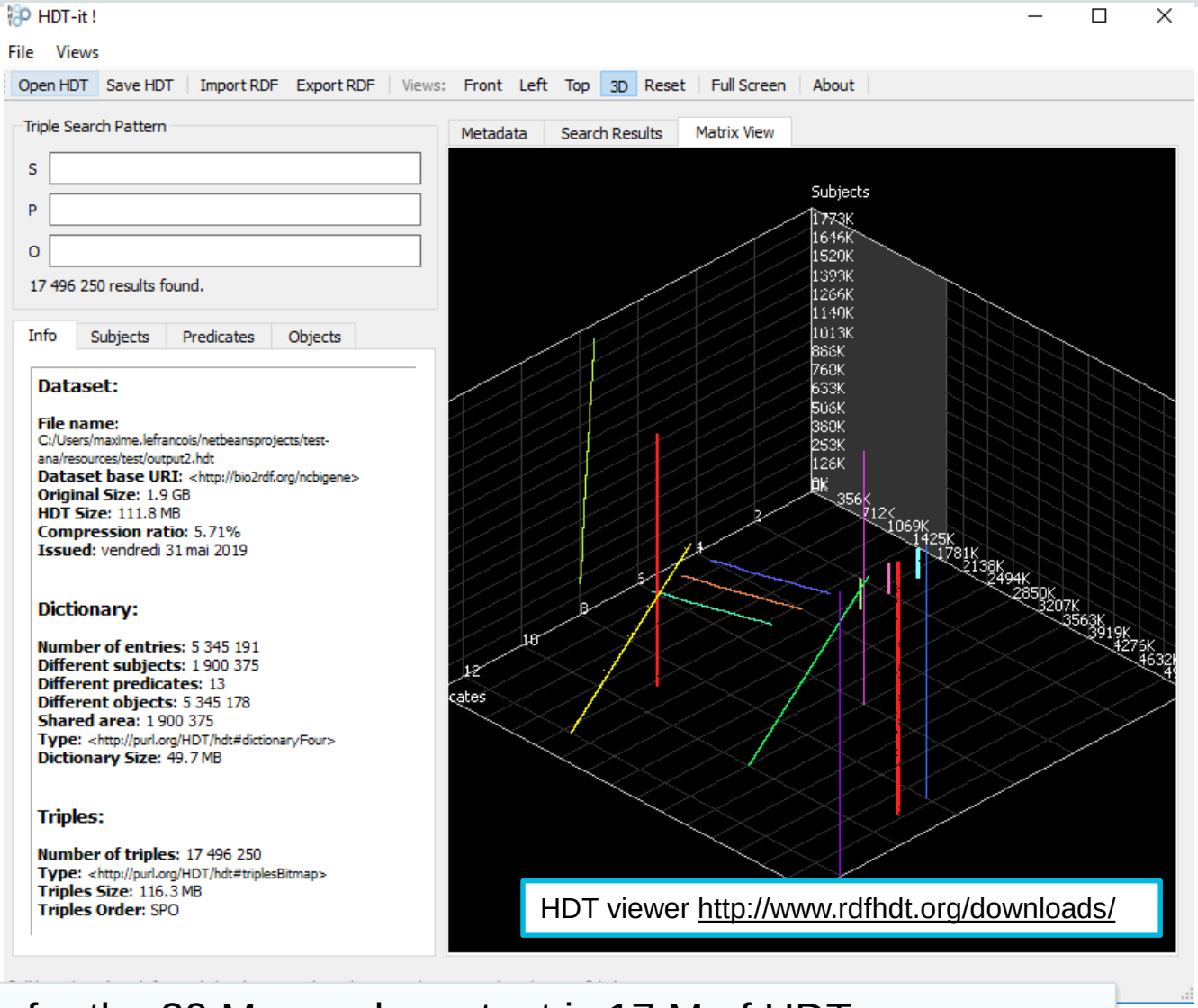
Additions in the past weeks

- Simple integration



Additions in the past weeks

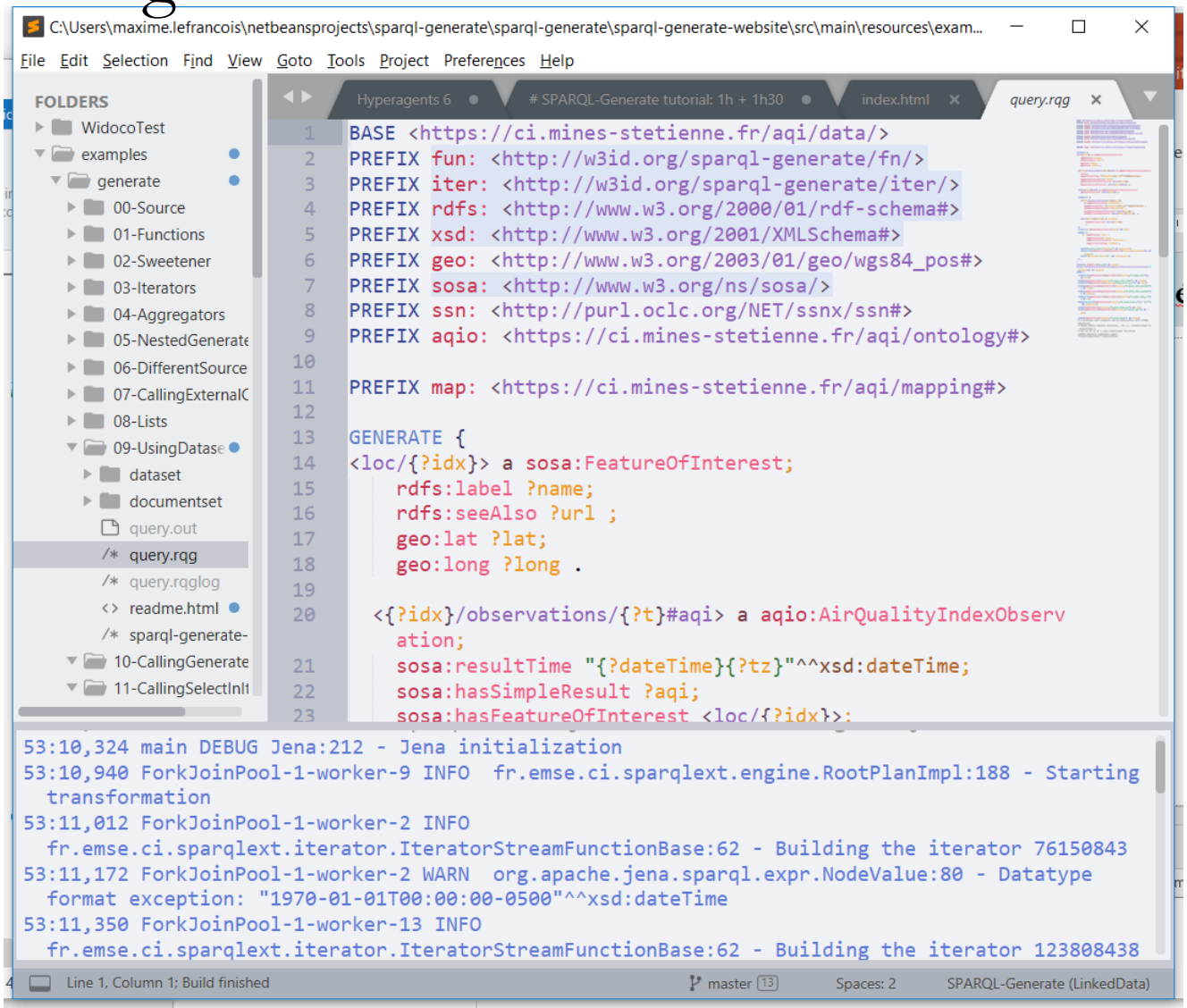
- HDT



20 seconds now only for the 20 M sample, output is 17 M of HDT
9 min, 20 sec for the 145 M file, output is the same "only" 114 M HDT

Additions in the past weeks

- Sublime Package



Institut Mines-Télécom

Take away message

1. RDF as lingua franca
2. Research in Ontology Engineering applied to engineering
3. Sem Web – based solutions for semantic interoperability

Maxime Lefrançois
Maxime.Lefrancois@emse.fr
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

Ontologies, langages, tools, to ease the usage of RDF as a lingua franca