Convering BIBFRAME to Schema.org

Osma Suominen

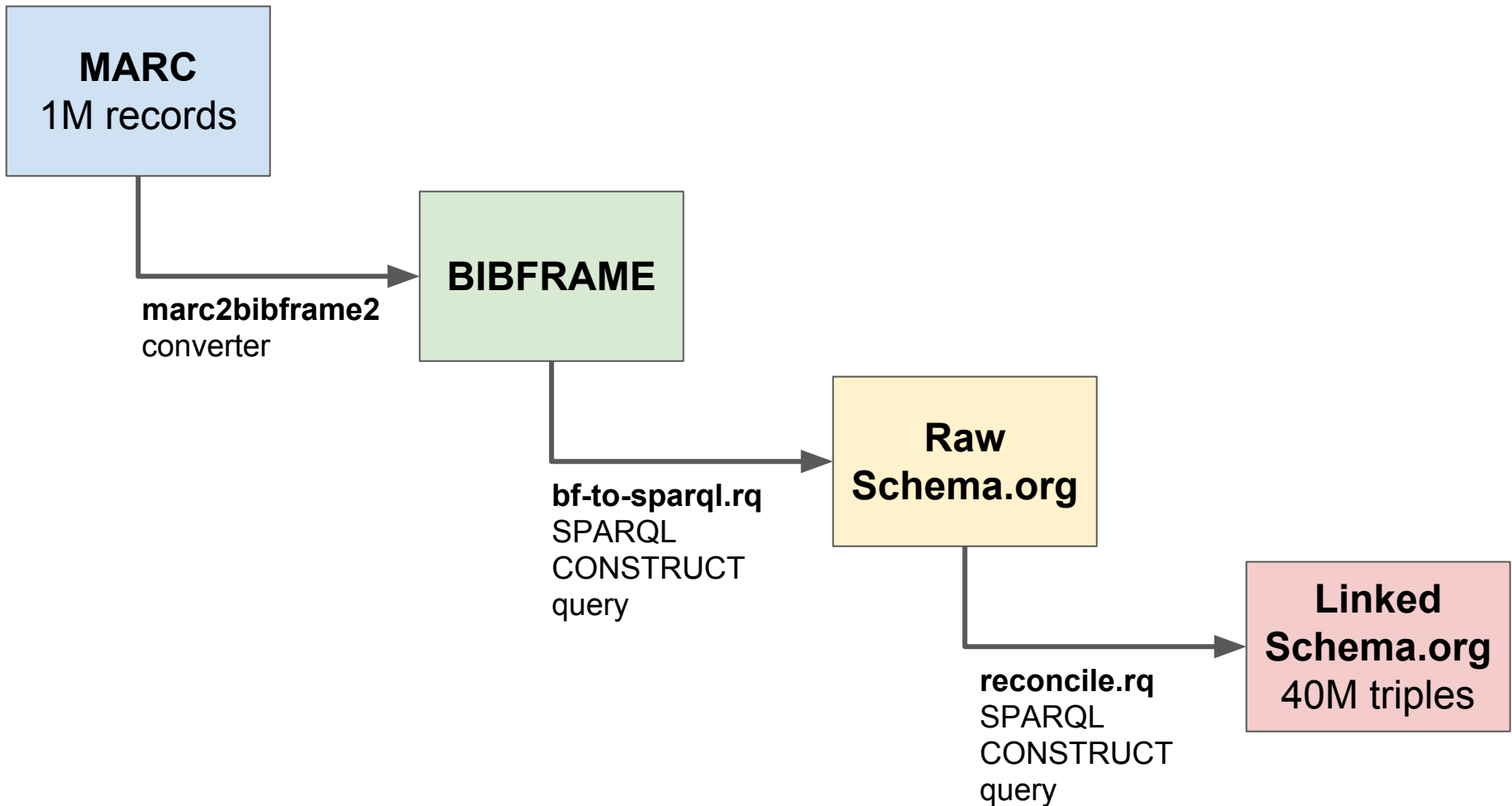European BIBFRAME Workshop

Florence, Italy
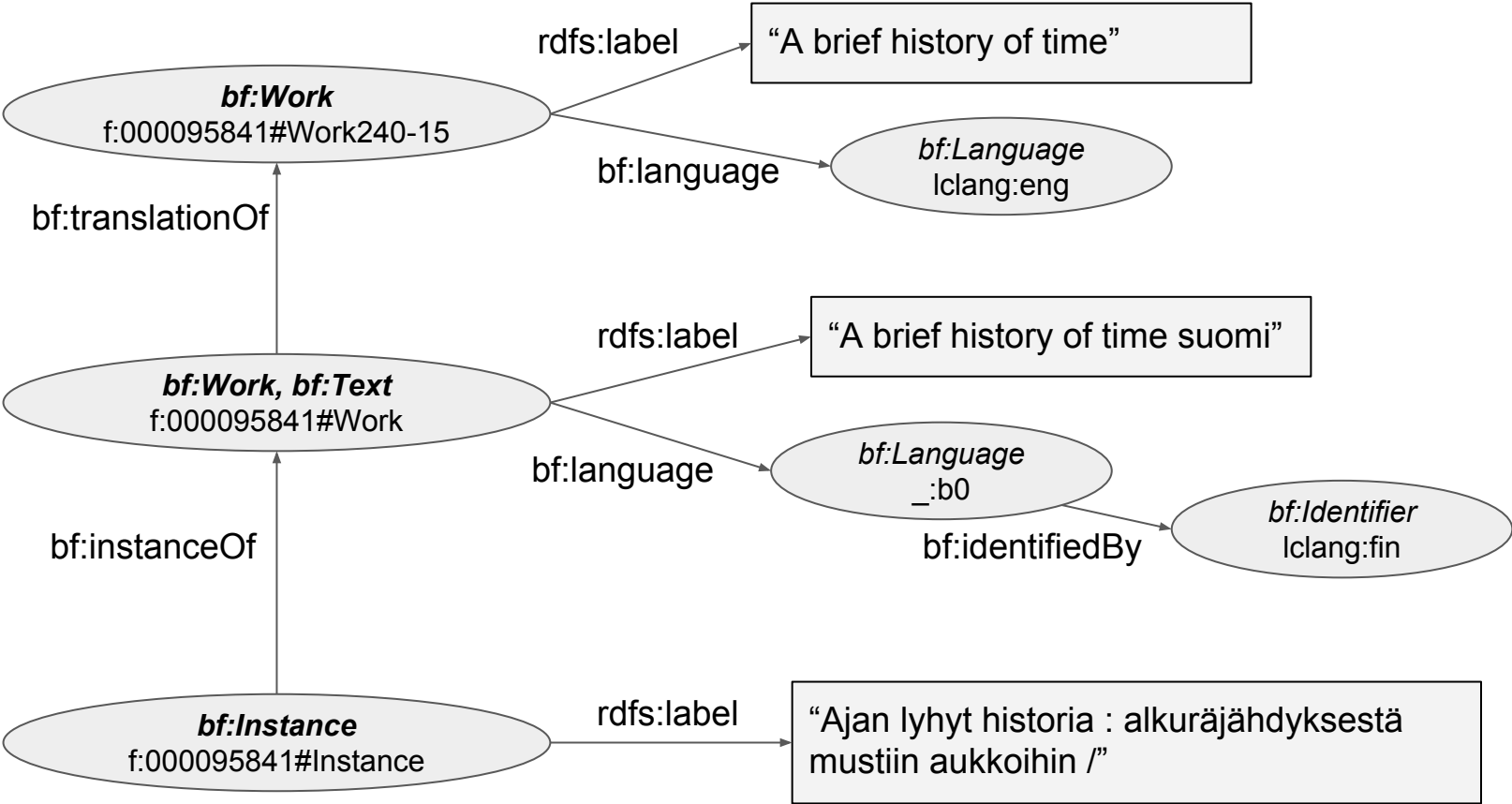
18 September 2018

Why?

1. Schema.org has become the *lingua franca* of the Web

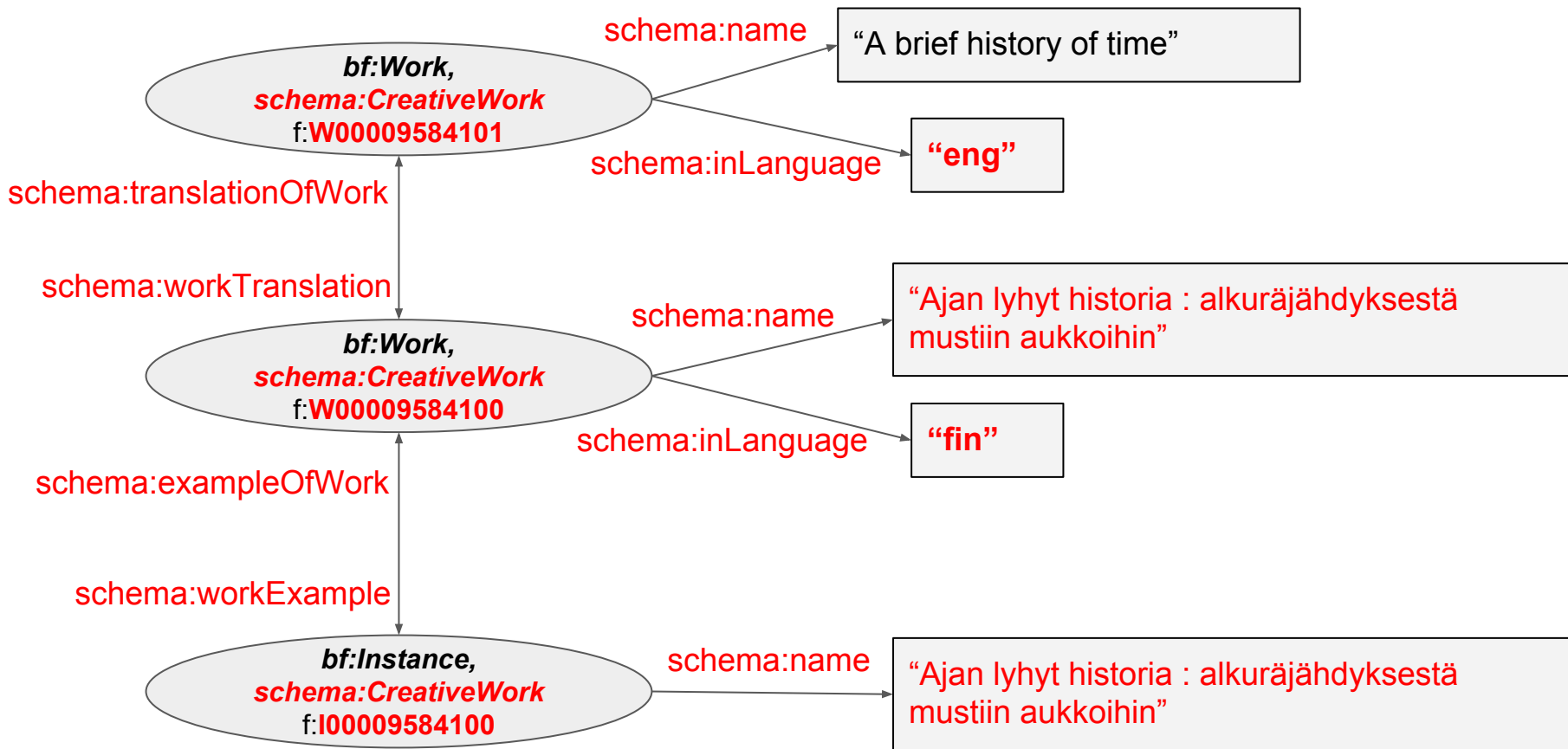2. Big picture view of our data: entities and their relationships

3. KISS principle

It is useful to go via BIBFRAME, but for our current use cases and experiments, Schema.org seems like a better fit

MARC
1M records
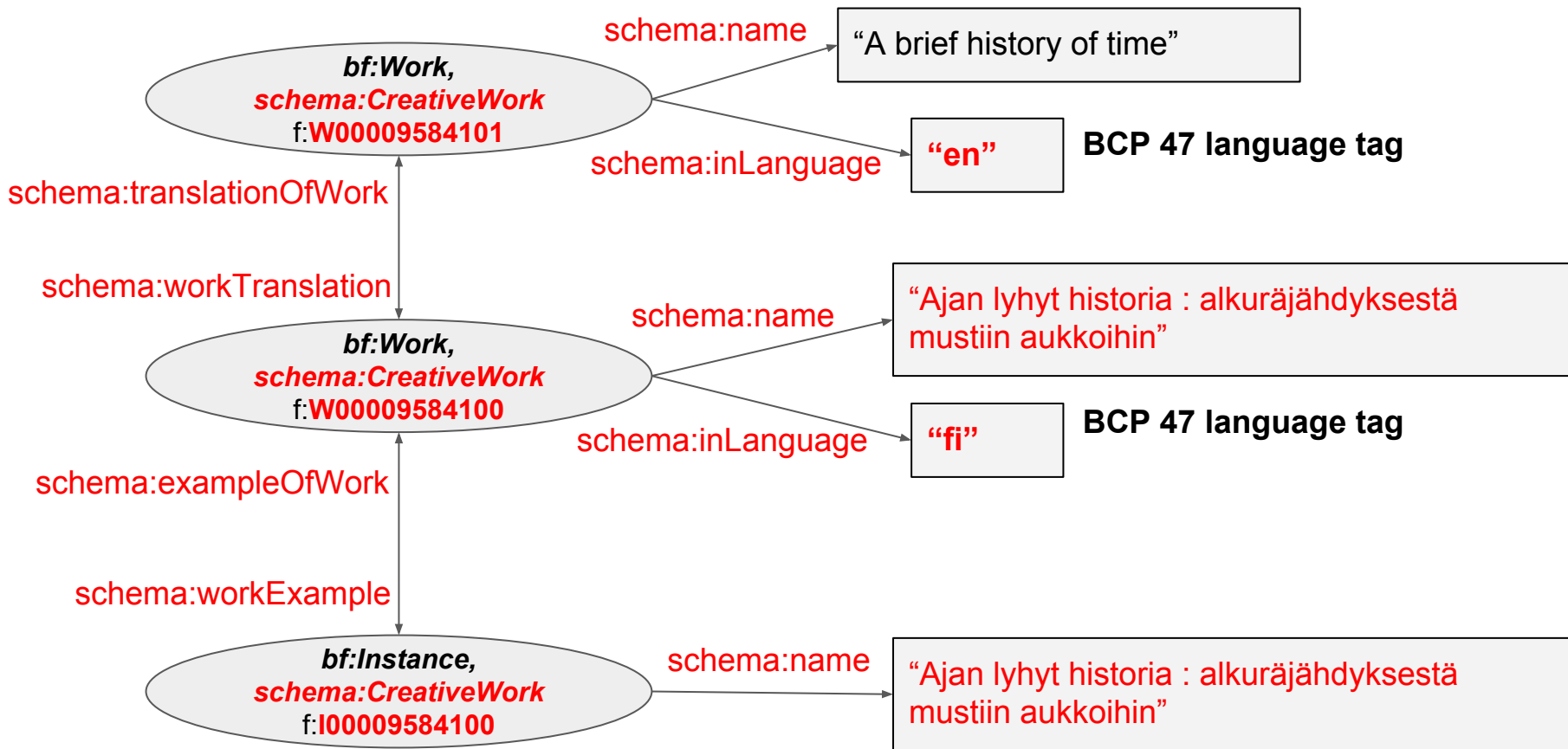
marc2bibframe2
converter

BIBFRAME

bf-to-sparql.rq
SPARQL
CONSTRUCT
query

**Raw
Schema.org**

reconcile.rq
SPARQL
CONSTRUCT
query

**Linked
Schema.org**
40M triples

BIBFRAME 2.0

Raw Schema.org

Reconciled Schema.org

[bf-to-sparql.rq](.) **query** (328 lines)

Observations and lessons learned

1. The conversion is relatively simple on a logical level, but **the devil is in the details**

2. The LOC marc2bibframe2 converter makes a good job of extracting information from a MARC record, but:
   a. many aspects are still just strings after conversion
   b. "one record at a time" conversion requires separate merging of works
   c. identity management is a real challenge, as the underlying MARC records may change at any time

Thank you!
Questions?

osma.suominen@helsinki.fi - @OsmaSuominen

http://data.nationallibrary.fi - @NatLibFiData

Code:
https://github.com/NatLibFi/bib-rdf-pipeline
https://github.com/NatLibFi/bib-lod-ui

These slides: **http://tinyurl.com/bf-to-schema**