RDF as a Universal Healthcare Exchange Language

David Booth, Ph.D.
KnowMED, Inc.

2013 Semantic Technology and Business Conference
San Francisco, CA
Latest version of these slides:
http://dbooth.org/2013/semtech/slides/03-DavidBooth-rdf-as-universal.pdf

Imagine a world

Imagine a world

in which **all healthcare systems**

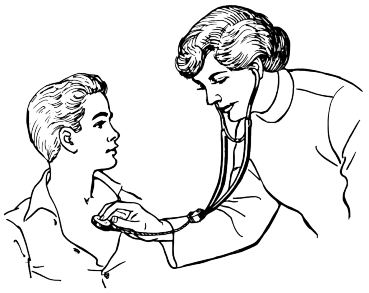
speak the **same language**

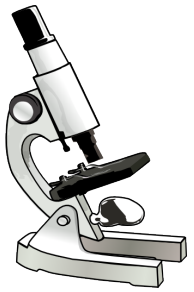
with the **same meanings**

covering **all healthcare.**

What would it be like?

- Better treatment
- Better research
- Lower cost





Goal: True semantic interoperability

Will RDF get us there?

- No. But it will get us closer.
- And along with the right policy incentives, RDF can get us *much* closer.

Why RDF?

Why RDF?

1. Semantics, not syntax
   - Syntax independent
   - Given RDF mappings, existing healthcare information formats can be viewed as RDF!

Why RDF?

1. Semantics, not syntax
2. Self describing
   - Concepts are identified by URIs
   - URIs can be dereferenceable to concept definitions
   - Helps bootstrap adoption of vocabularies

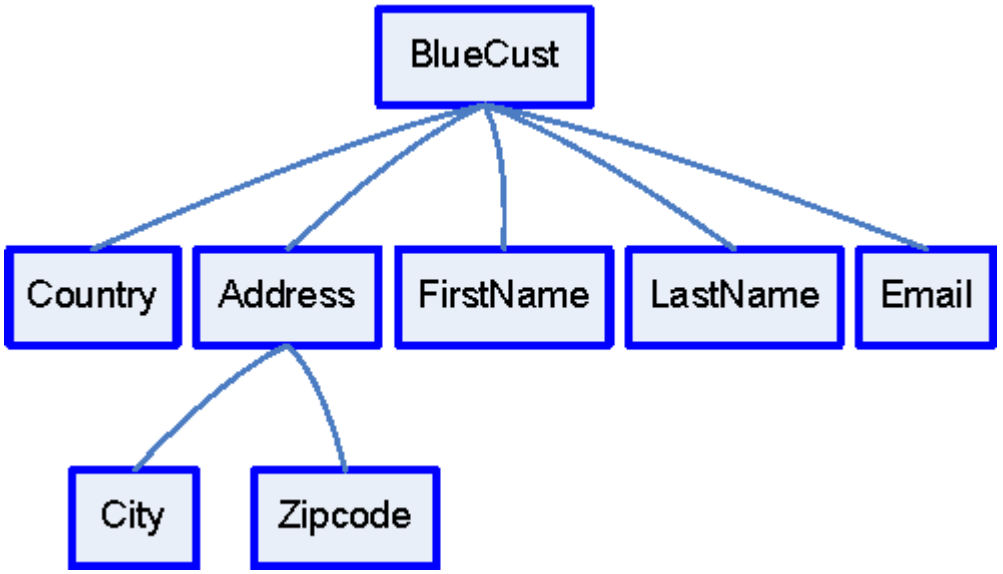
Why RDF?

1. Semantics, not syntax
2. Self describing
3. Schema promiscuous
   - Multiple data models peacefully co-exist
     - Semantically linked
     - In the same data
   - Unlike schema-centric languages, e.g. XML

Why RDF?

Schema promiscuous

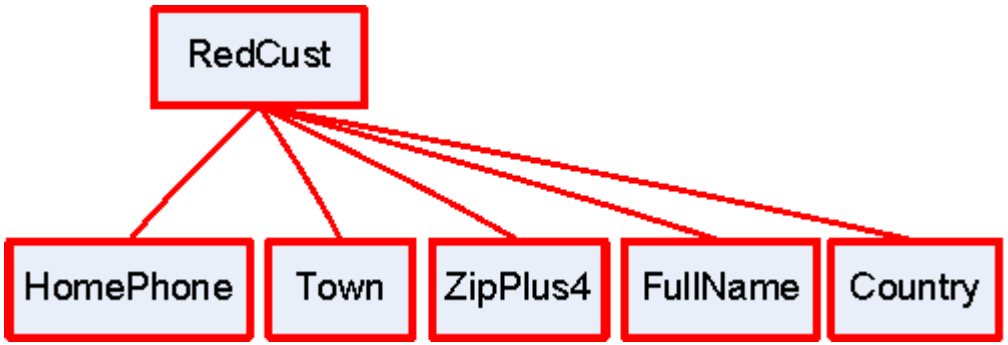
- Blue App has model



Why RDF?

Schema promiscuous

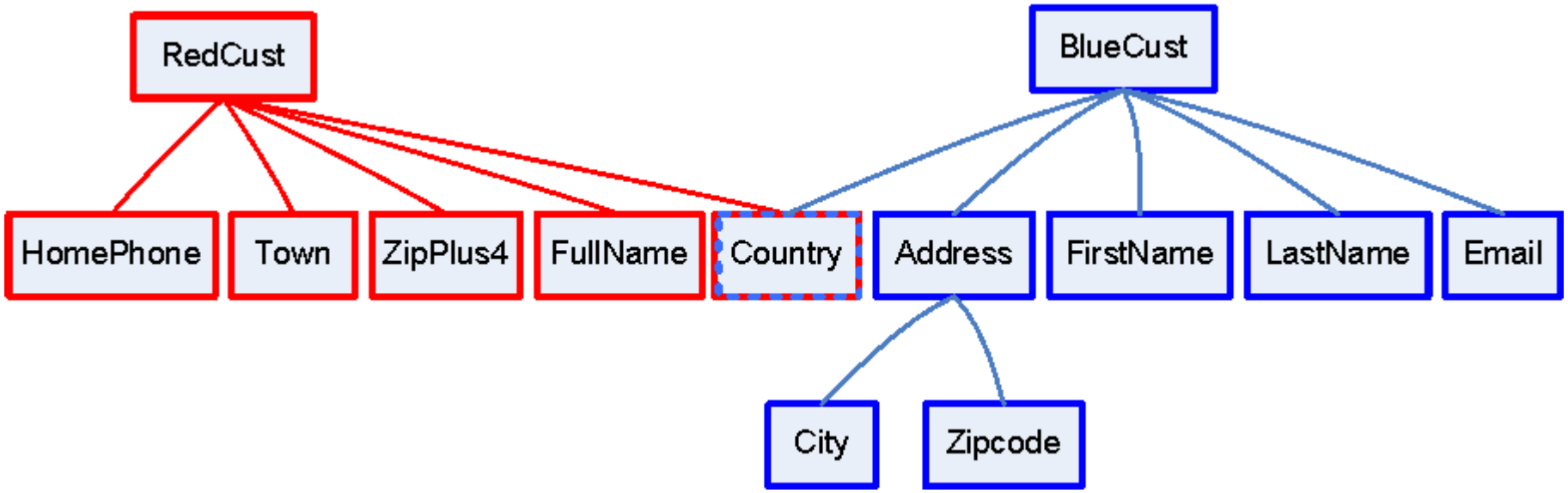
- Red App has model



Why RDF?

Schema promiscuous

- Merge RDF data
- Same nodes (URIs) join automatically

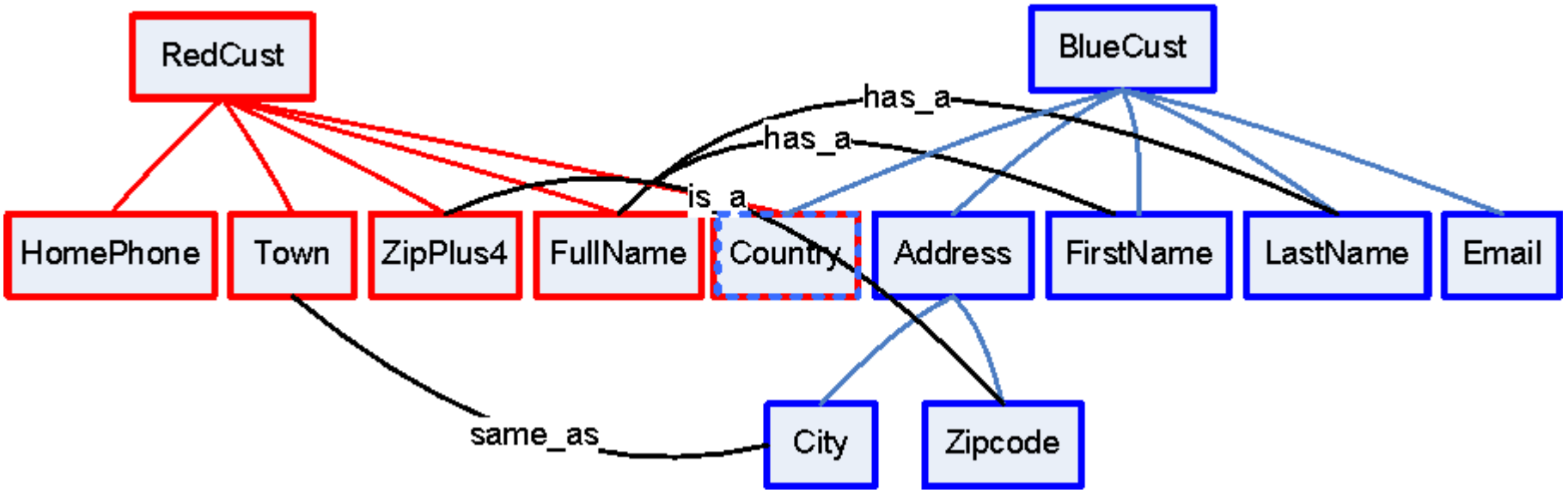


Multiple models peacefully coexist

Why RDF?

Schema promiscuous

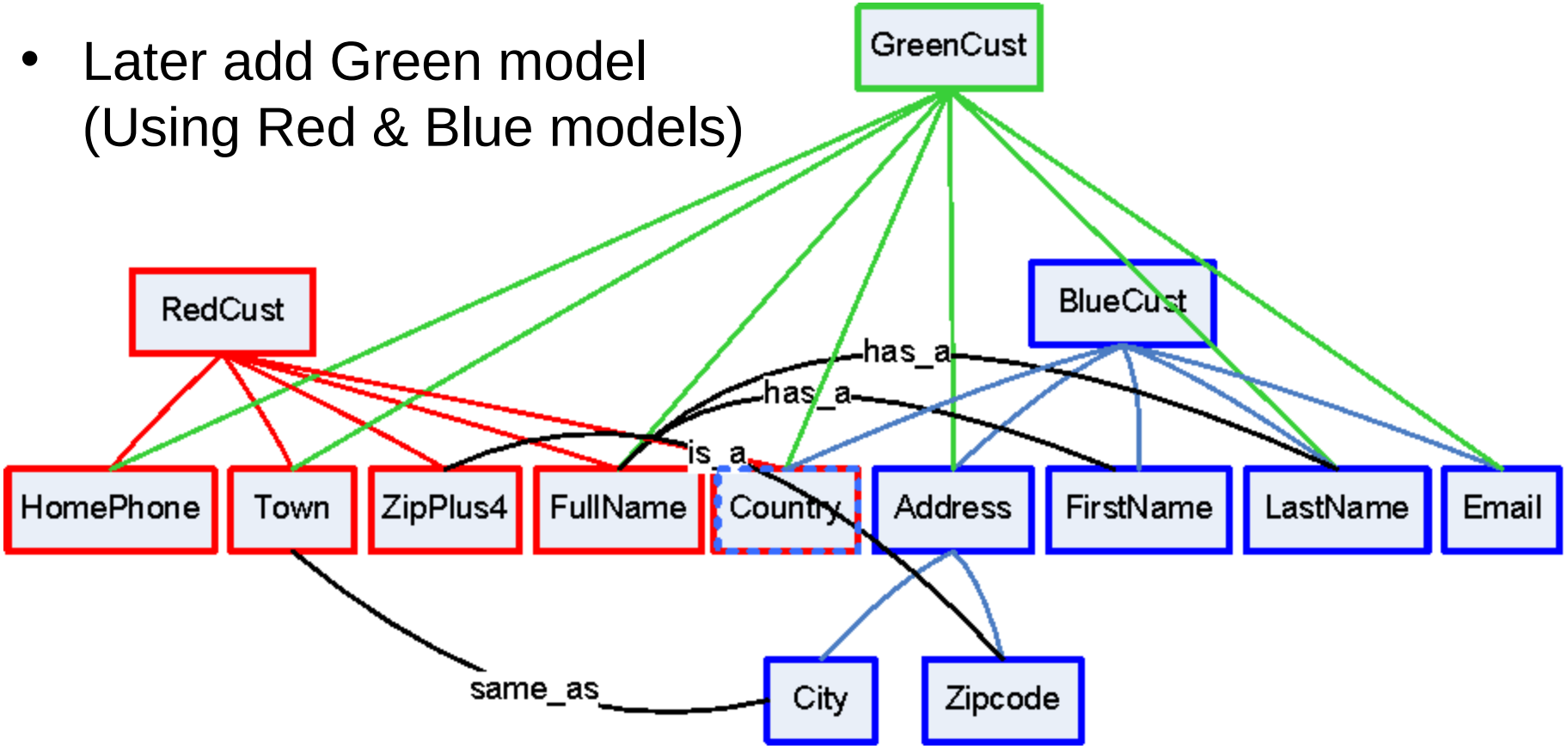
- Add relationships and rules
- (Relationships are also RDF)



Why RDF?

Schema promiscuous

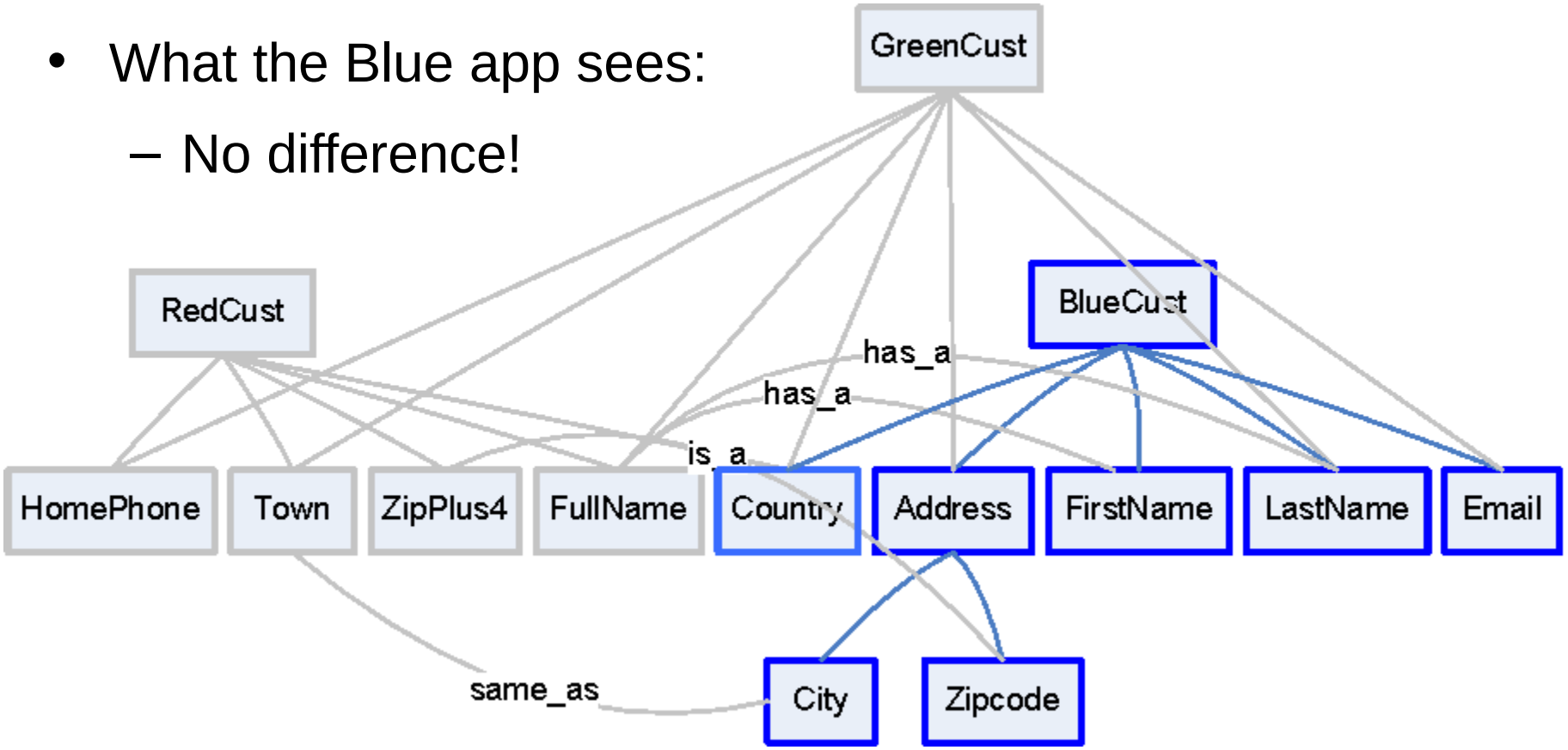
- Later add Green model (Using Red & Blue models)



Why RDF?

Schema promiscuous

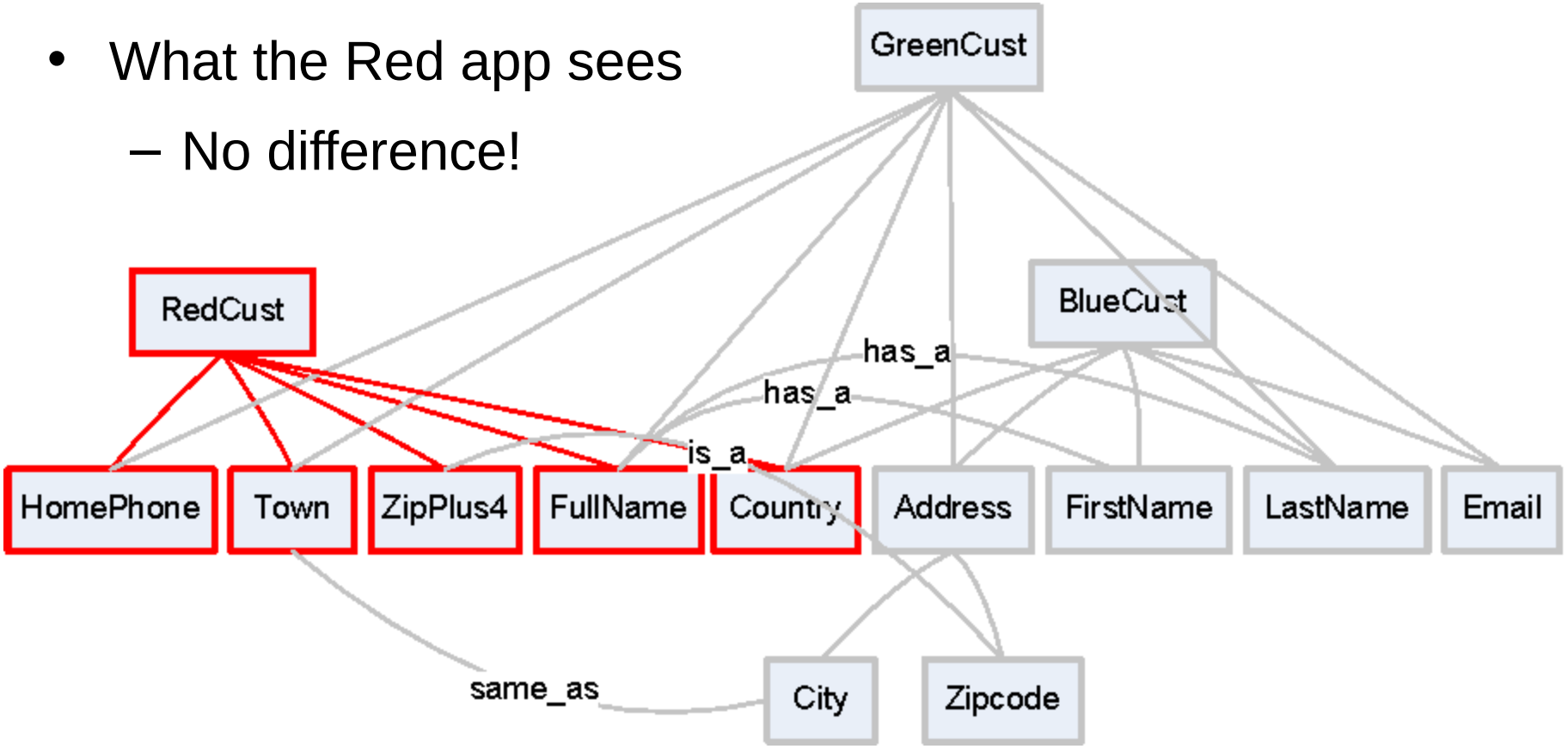
- What the Blue app sees:
  - No difference!



Why RDF?

Schema promiscuous

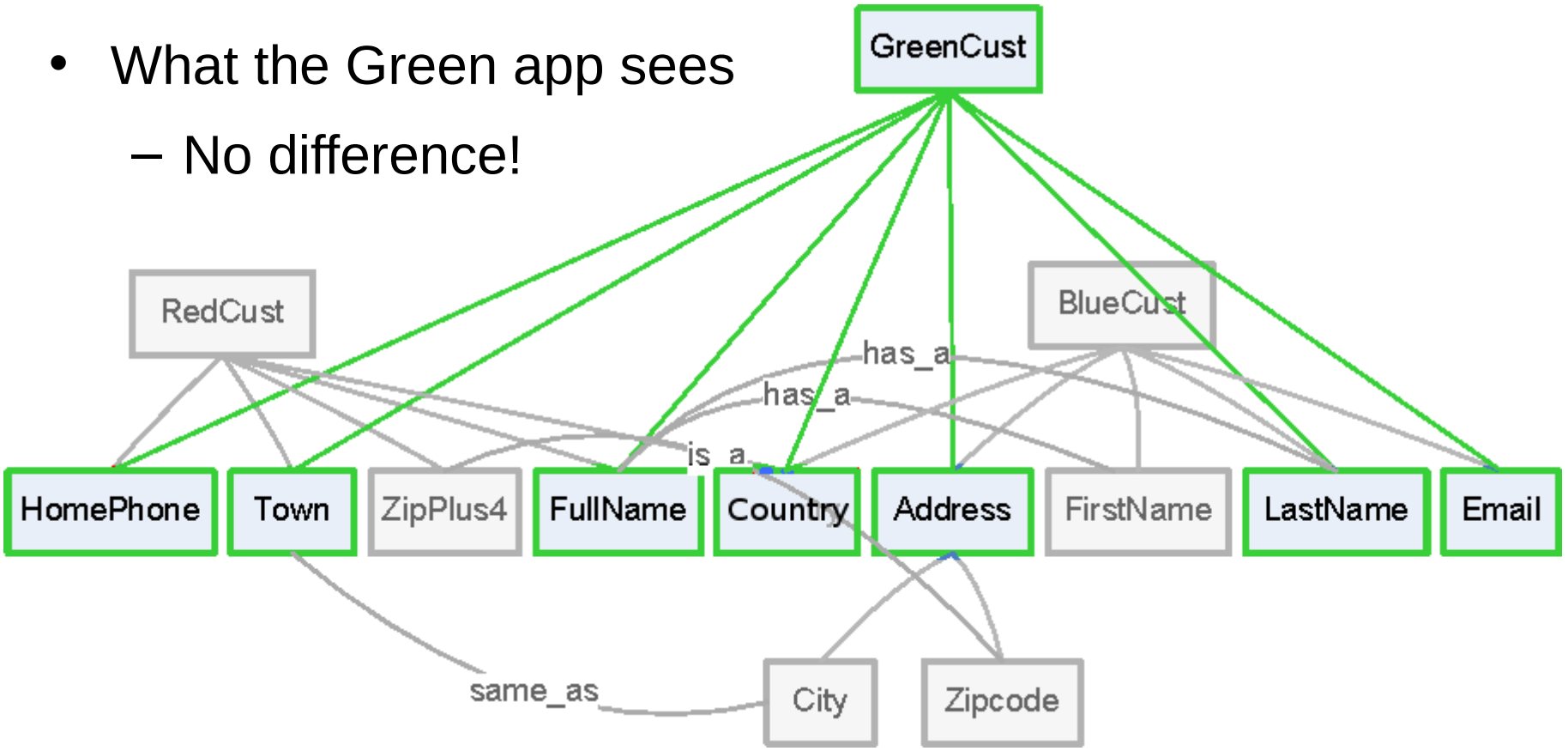
- What the Red app sees
  - No difference!



Why RDF?

Schema promiscuous

- What the Green app sees
  - No difference!



Why RDF?

1. Semantics, not syntax
2. Self describing
3. Schema promiscuous
4. **Neutral, mature, international standard**

Why RDF?

1. Semantics, not syntax
2. Self describing
3. Schema promiscuous
4. Neutral, mature, international standard

Best available option for a universal healthcare exchange language!

"RDF as a Universal Healthcare Exchange Language": What does it mean?

Misconceptions

- Change EHR databases to RDF stores?
  - *No. Just transform to/from RDF for exchange.*
- Discard existing healthcare information standards (HL7, SNOMED, LOINC, etc.)?
  - *No. Leverage them by mapping to RDF.*

What does it mean?

1. Use RDF as a *substrate* for exchange of healthcare information

- Exchange data either in:
  - a generic RDF syntax; or
  - a common format that can be mapped to the RDF model
- E.g. Turtle or HL7

What does it mean?

1. Use RDF as a *substrate* for exchange of healthcare information

2. Adopt *standard syntactic mappings* of common healthcare information formats to the RDF model

- RDF is *syntax independent*
- Given a mapping to the RDF model, *any* format can be viewed as RDF
- Some RDF mappings have already been created

What does it mean?

1. Use RDF as a *substrate* for exchange of healthcare information

2. Adopt *standard syntactic mappings* of common healthcare information formats to the RDF model

3. Adopt *standard, self-describing URIs* for healthcare concepts

- Dereferenceable to free and open definitions
- For all common vocabularies: SNOMED, LOINC, etc.
- Also for people, places and institutions
- NOTE: *Definitions* would be public; values would not

What does it mean?

1. Use RDF as a *substrate* for exchange of healthcare information
2. Adopt *standard syntactic mappings* of common healthcare information formats to the RDF model
3. Adopt *standard, self-describing URIs* for healthcare concepts

4. Adopt *standard semantic mappings* between overlapping concepts

- Both within and between vocabularies
- E.g., v1:SystolicBP subsumes v2:BP_Systolic_Sitting

What does it mean?

1. Use RDF as a *substrate* for exchange of healthcare information
2. Adopt *standard syntactic mappings* of common healthcare information formats to the RDF model
3. Adopt *standard, self-describing URIs* for healthcare concepts
4. Adopt *standard semantic mappings* between overlapping concepts

Also helpful, but beyond the scope of this workshop:
[5. Use RESTful Linked Data principles]

RDF as a Universal Healthcare Exchange Language

1. Use RDF as a *substrate* for exchange of healthcare information

2. Adopt *standard syntactic mappings* of common healthcare information formats to the RDF model

3. Adopt *standard, self-describing URIs* for healthcare concepts

4. Adopt *standard semantic mappings* between overlapping concepts

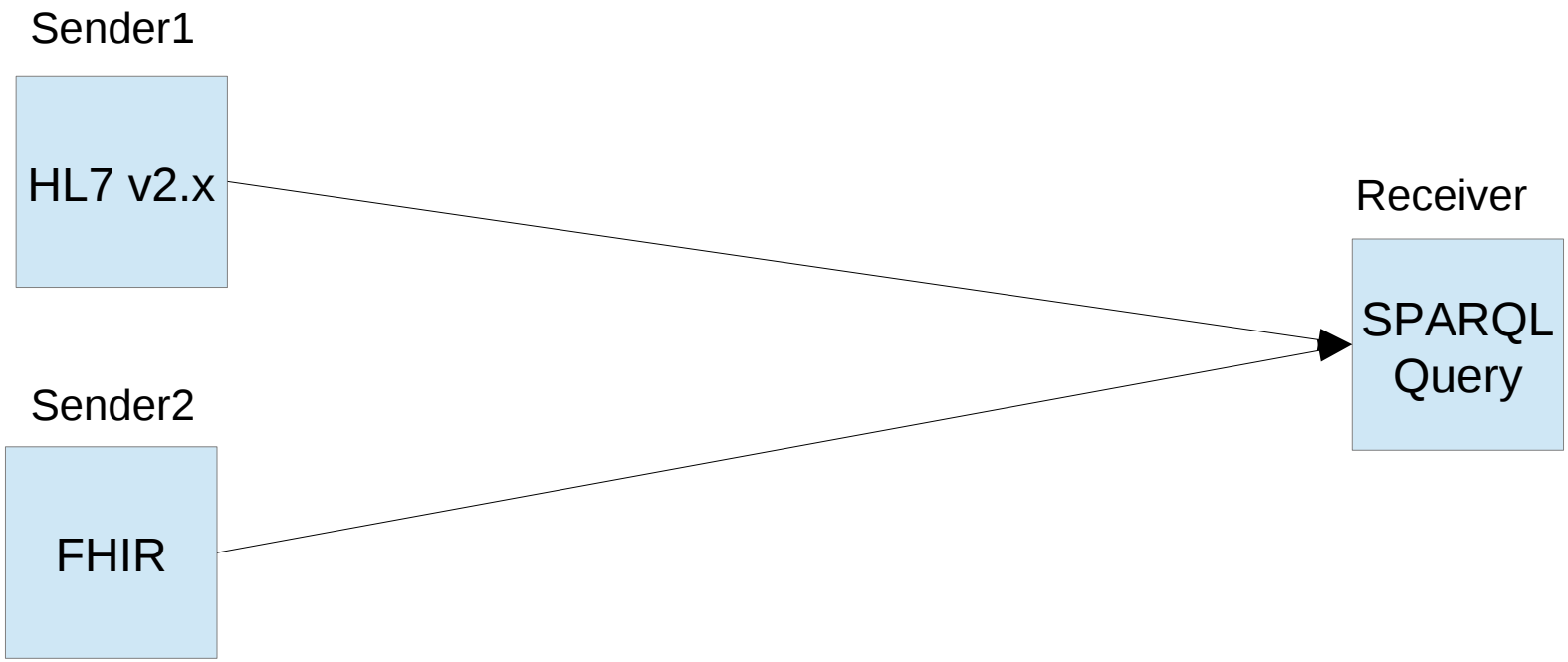
Also helpful, but beyond the scope of this workshop:
[5. Use RESTful Linked Data principles]

Achieving adoption

- Bad news:
  - Healthcare players have no financial incentive to make data interoperable
- Good news:
  - Government agencies can create incentives
    - Either carrots or sticks

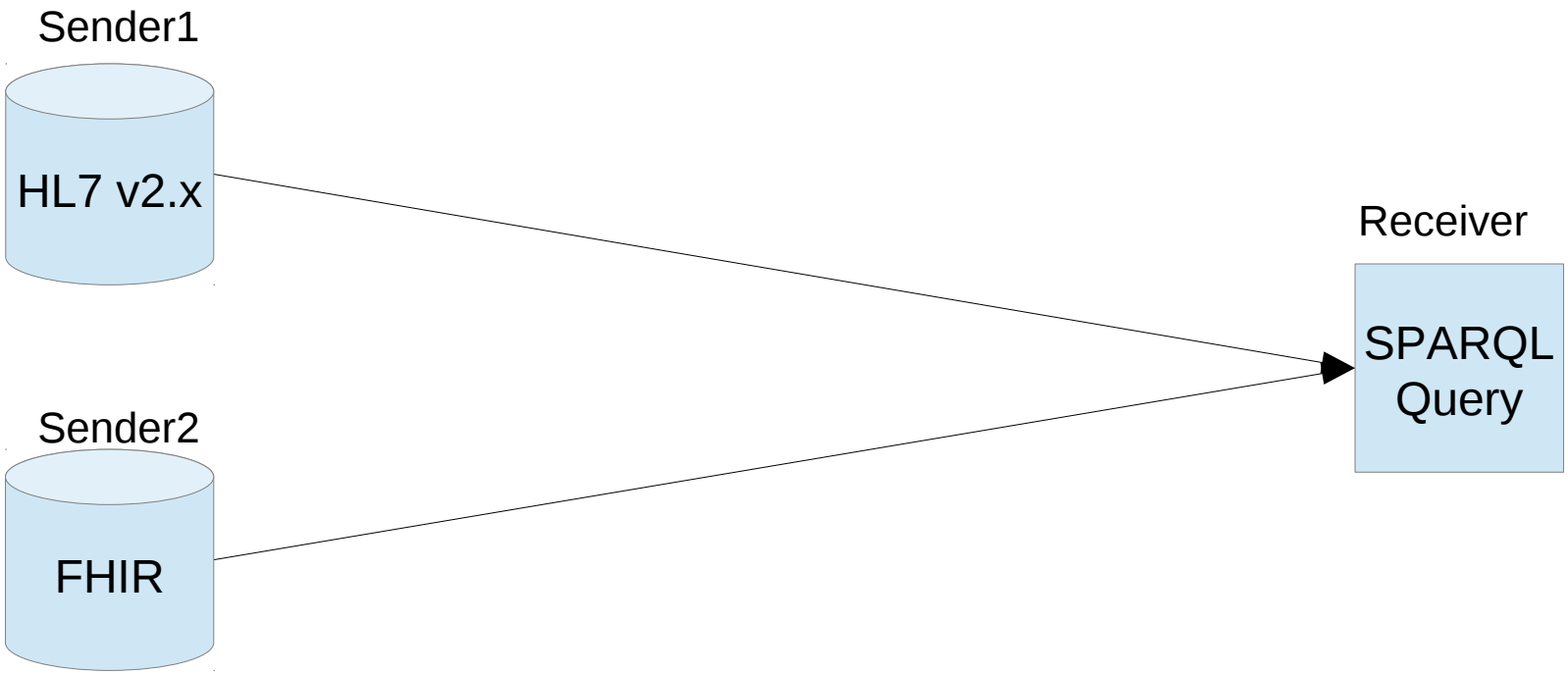
Government agencies must incentivize healthcare data interoperability!

Goal: Semantic Interoperability



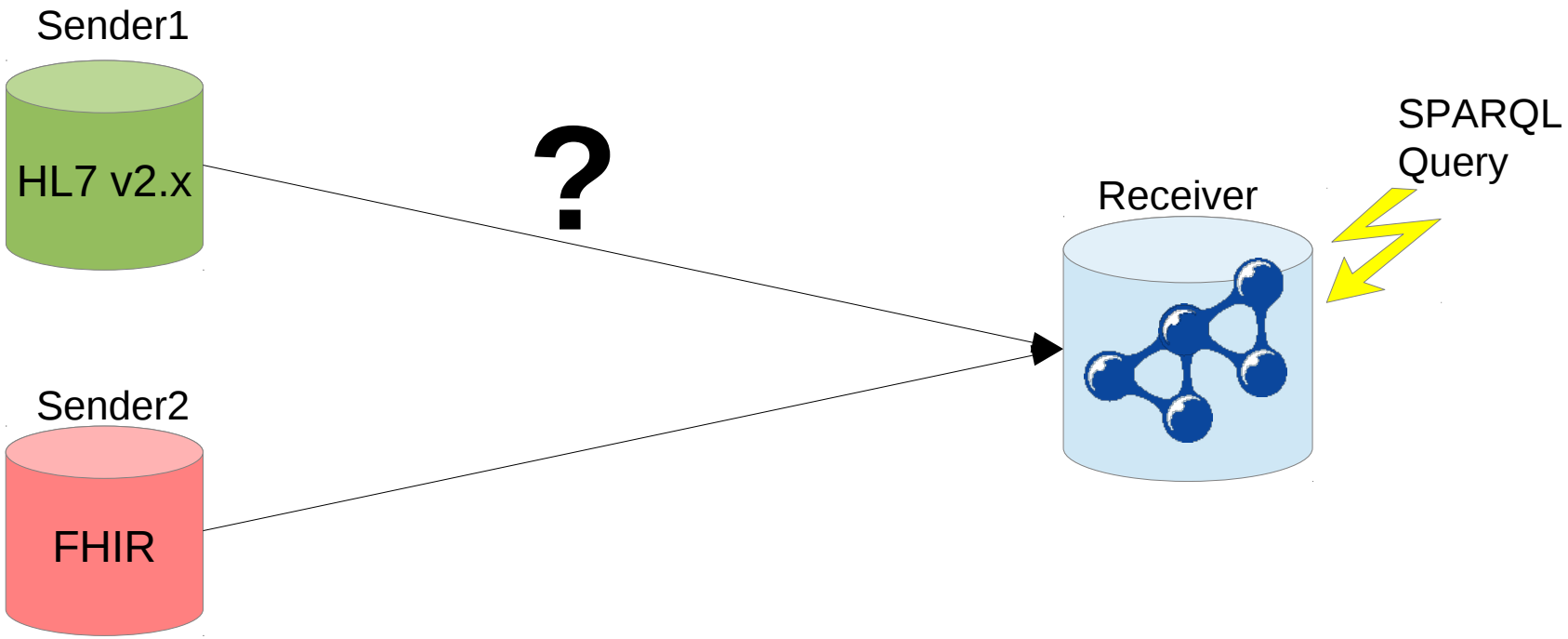
- Receiver wishes to combine data from Sender1 and Sender2

Step 1: Syntactic transformation



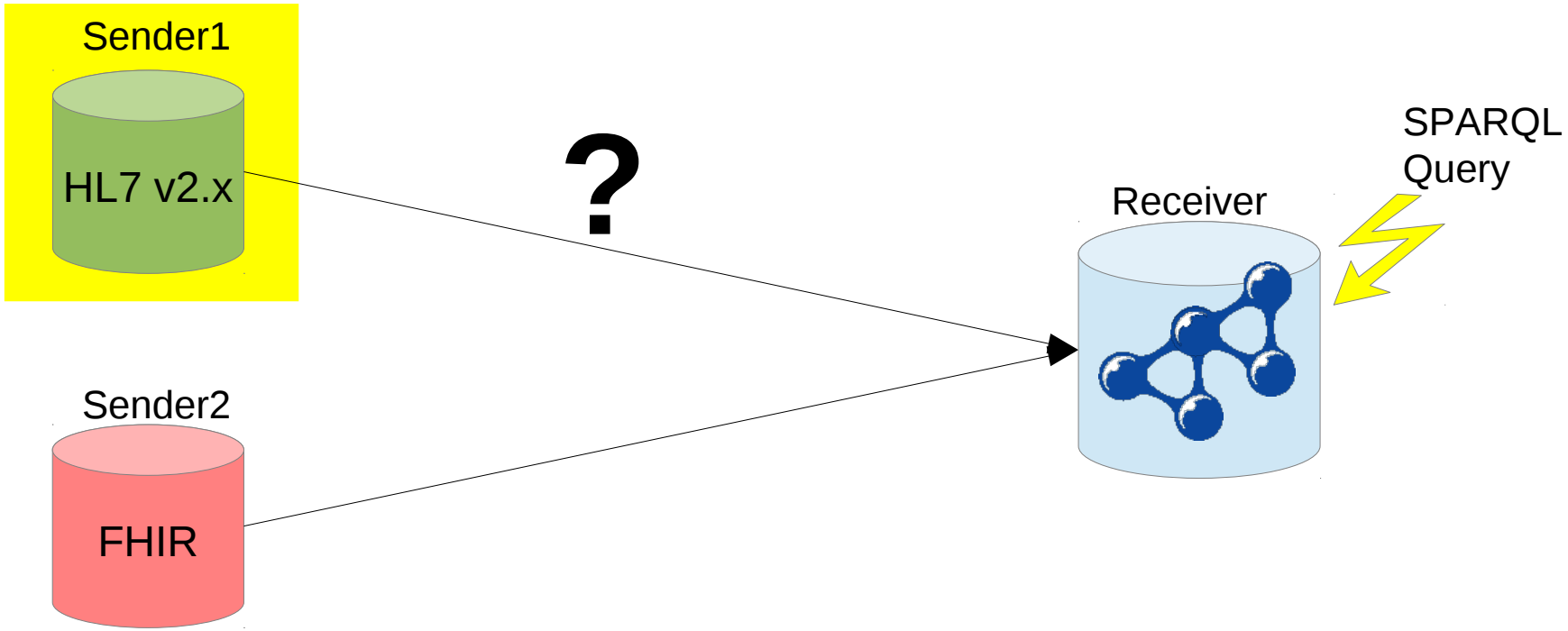
- Senders' native formats must be transformed to a common "substrate model"
- E.g., transform to RDF

Goal: Semantic interoperability



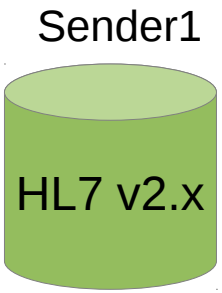
- Receiver wishes to query combined data
- But parties use different formats & vocabularies

Goal: Semantic interoperability



- Receiver wishes to query combined data
- But parties use different formats & vocabularies

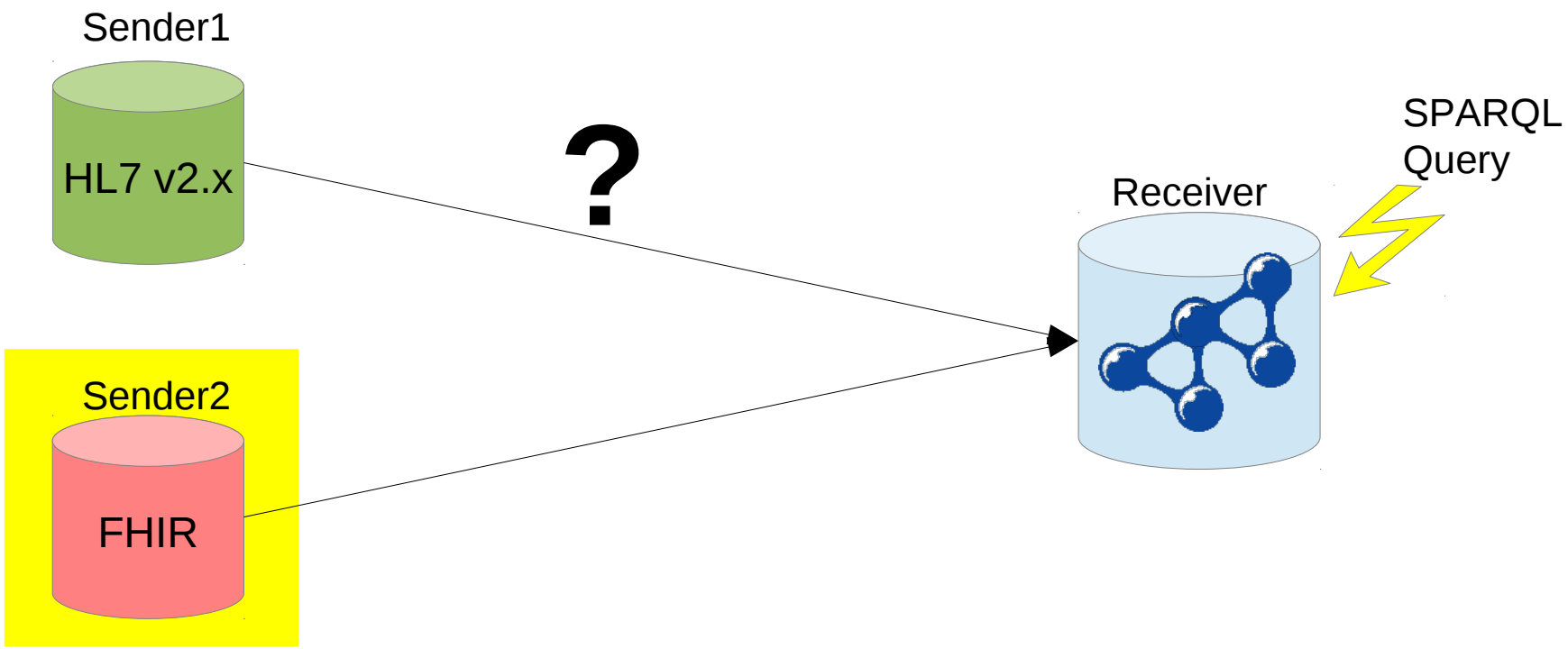
Sender1 data: HL7 v2.x




```
OBX|1|CE|3727-0^BPsystolic,
sitting||120||mmHg|
```

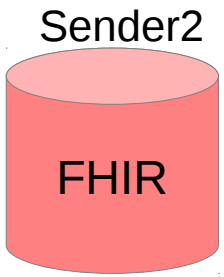
(Fictitious examples for illustration)

Goal: Semantic interoperability



- Now look at Sender2's data . . .

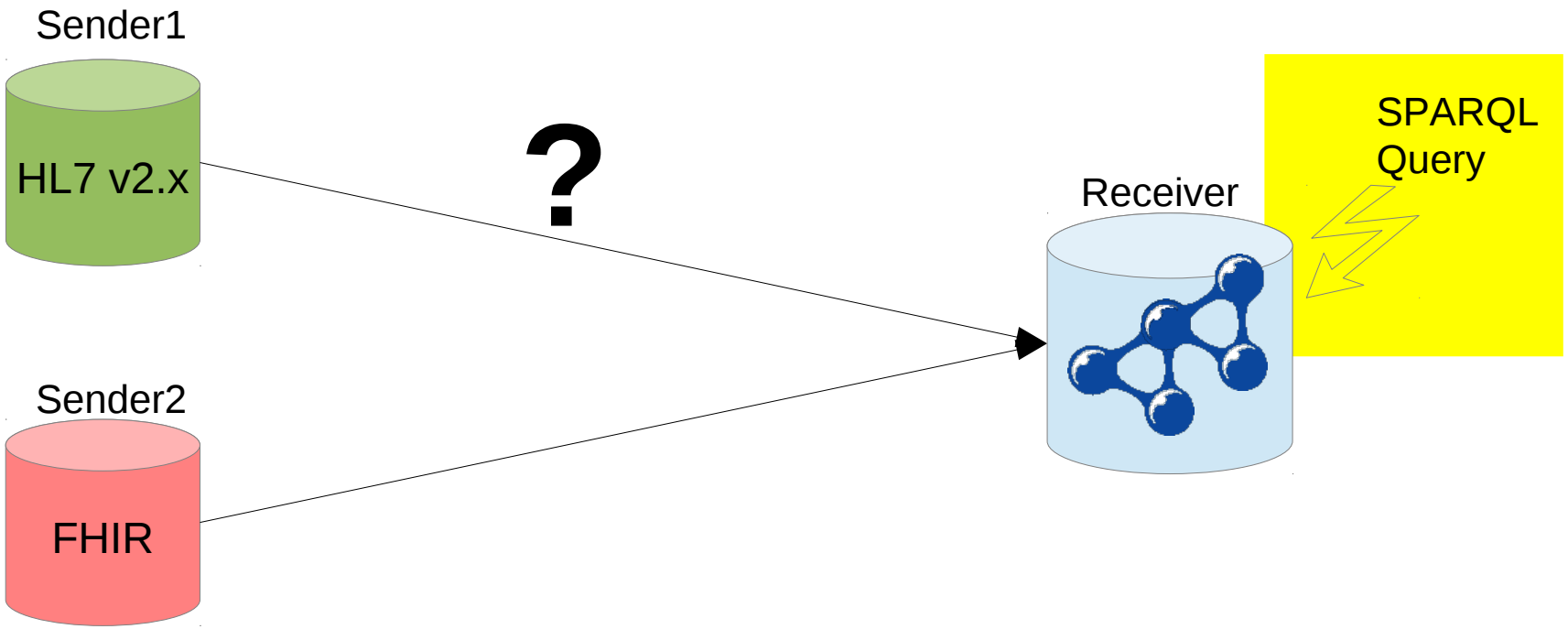
Sender2 data: FHIR




```
<Observation
      xmlns="http://hl7.org/fhir">
   <system value="http://loinc.org"/>
   <code value="8580-6"/>
   <display value="Systolic BP"/>
   <value value="107"/>
   <units value="mm[Hg]"/>
</Observation>
```

(Fictitious example for illustration)

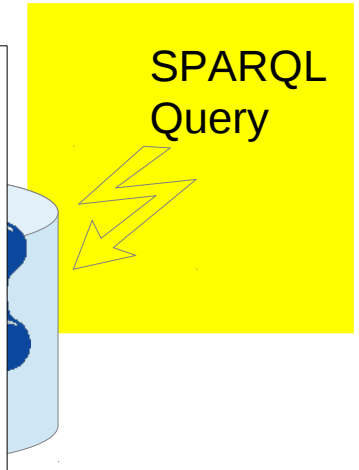
Goal: Semantic interoperability



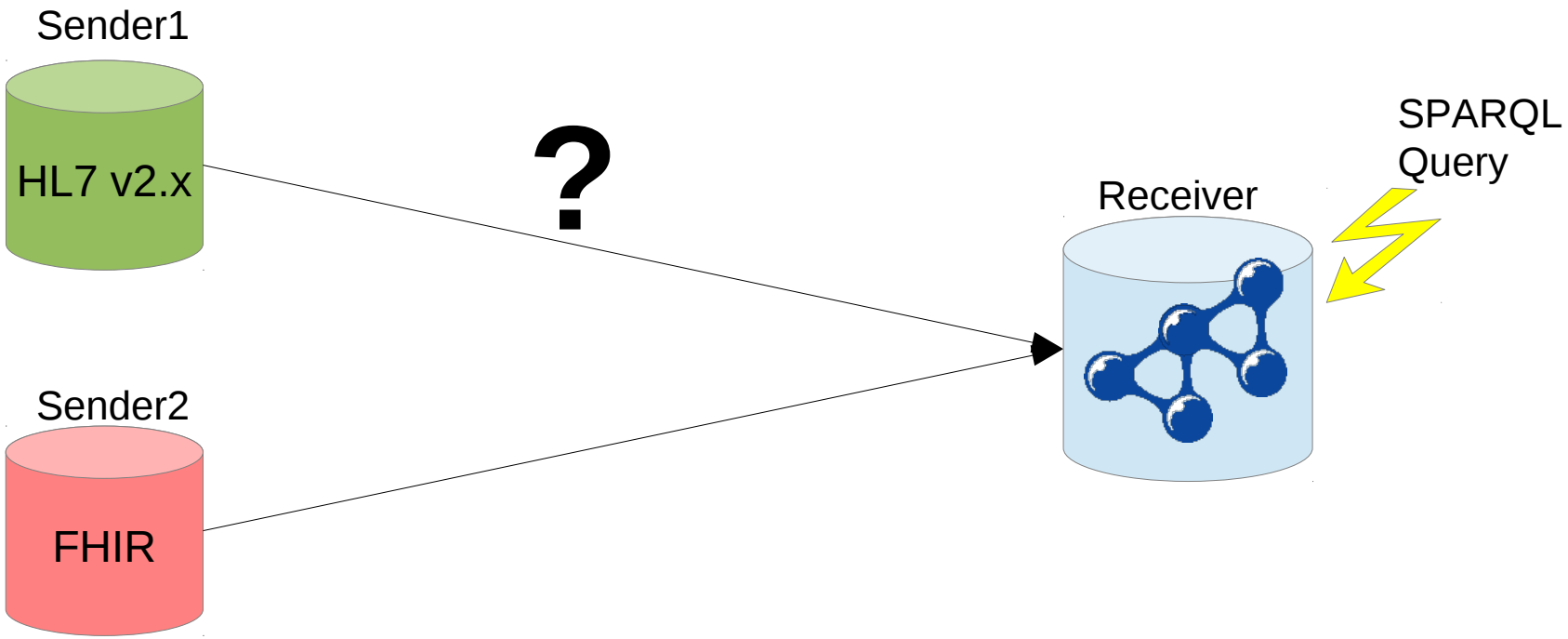
- Now look at Receiver's query . . .

Receiver SPARQL query

```
SELECT ?systolic
WHERE {
  ?observation a mOut:Observation ;
    a mOut:BP_systolic ;
    mOut:value ?systolic ;
    mOut:units mOut:mmHg . }
```

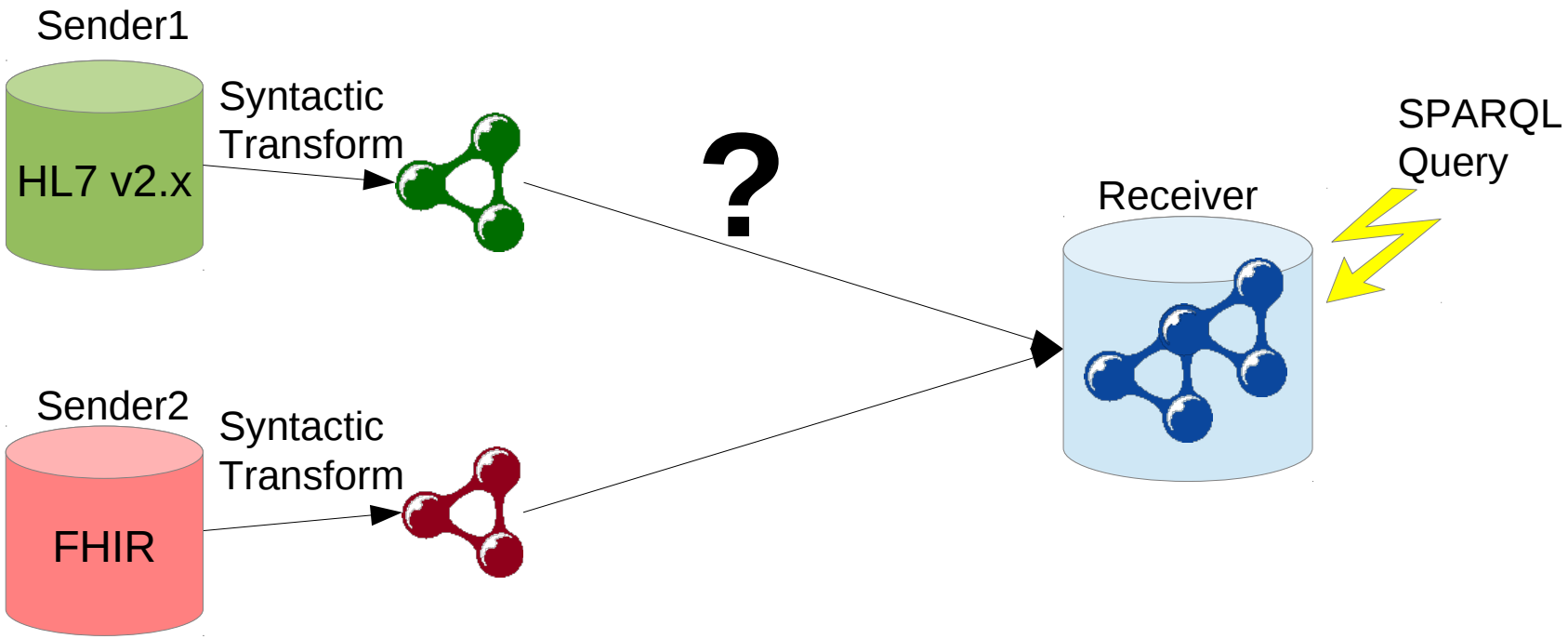


Goal: Semantic interoperability



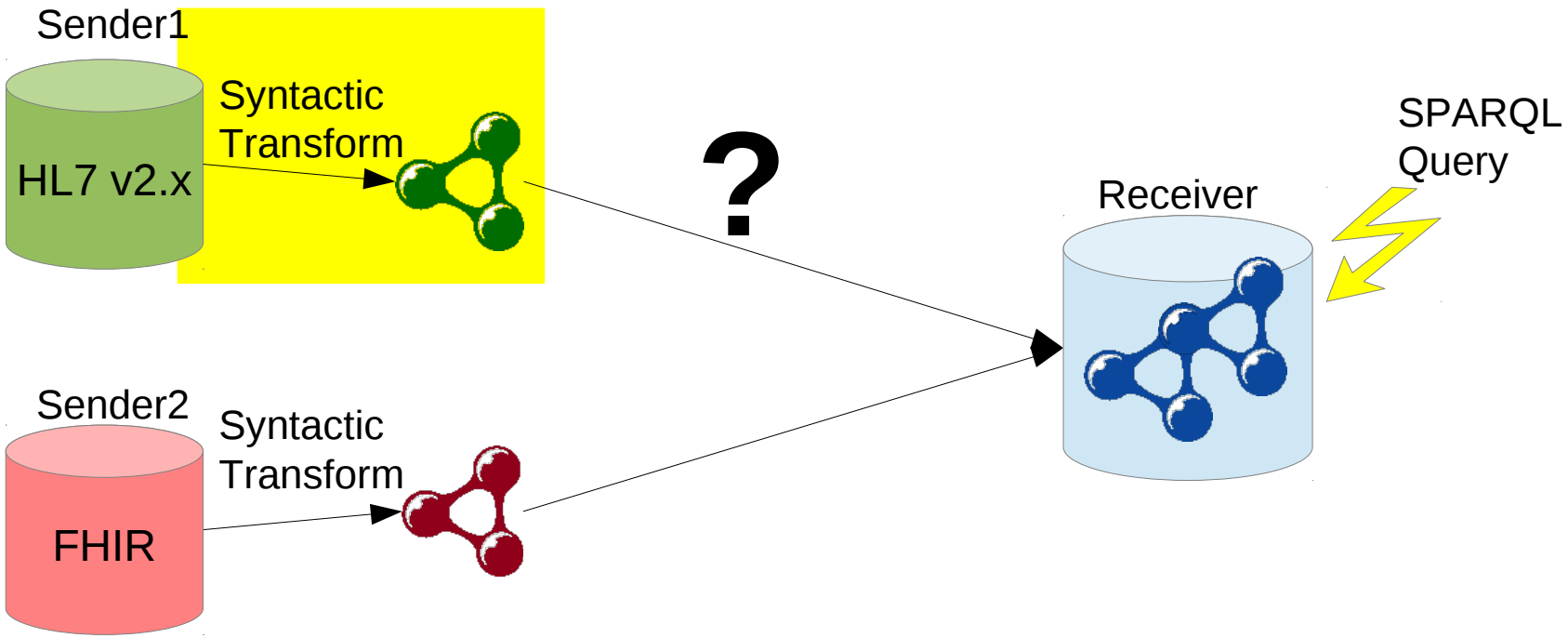
- What to do?

Step 1: Syntactic transformation



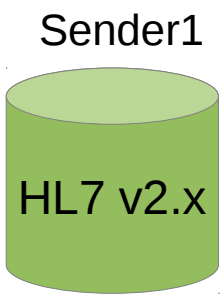
- Transform from source format to substrate model
- Allows data to be merged
- Data may not link semantically due to differing vocabularies

Step 1: Syntactic transformation



- Look at Sender1 RDF . . .

Sender1 syntactic transformation

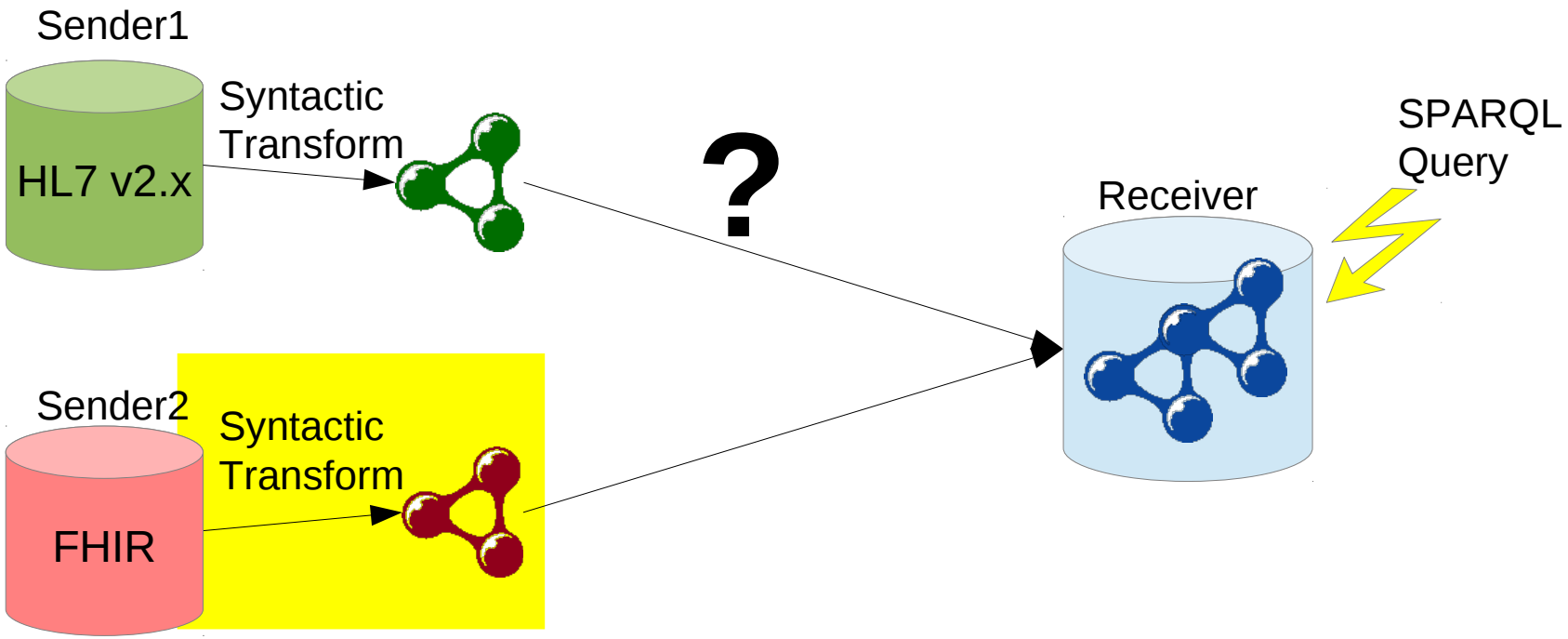



```
OBX|1|CE|3727-0^BPsystolic,
sitting||120||mmHg|
```

RDF

```
d1:obs042 a m1:PatientObservation ;
  m1:code "3727-0" ;
  m1:description "BPsystolic, sitting" ;
  m1:value 120 ;
  m1:units "mmHg" .
```

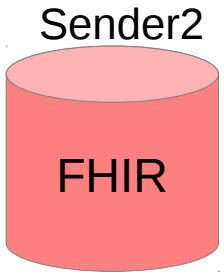
Step 1: Syntactic transformation



- Look at Sender2 RDF . . .

Sender2 syntactic transformation

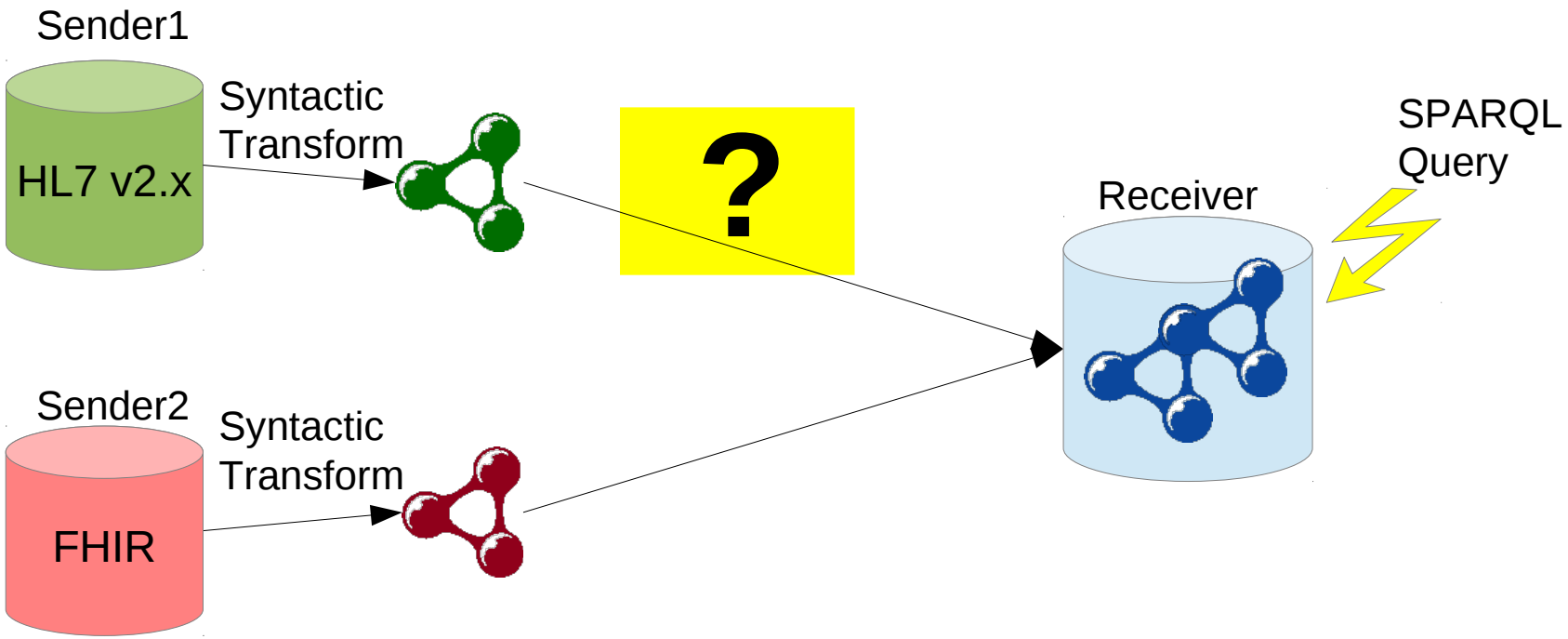
```
<Observation
      xmlns="http://hl7.org/fhir">
   <system value="http://loinc.org"/>
   <code value="8580-6"/>
   <display value="Systolic BP"/>
   <value value="107"/>
   <units value="mm[Hg]"/>
</Observation>
```



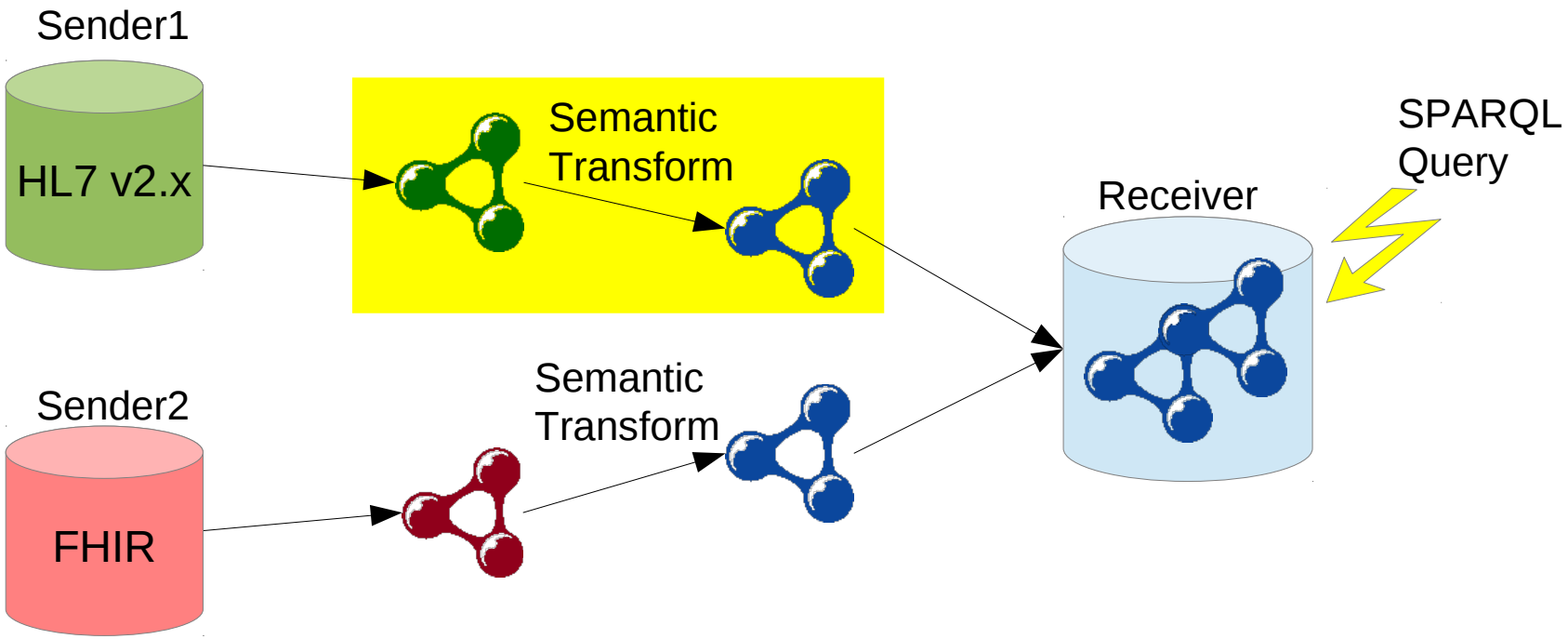
RDF

```
d2:obs-091 a m2:Observation ;
  m2:system "http://loinc.org" ;
  m2:code "8580-6" ;
  m2:display "Systolic BP" ;
  m2:value 107 ;
  m2:units "mm[Hg]" .
```

Step 1: Syntactic transformation



Step 2: Semantic transformation



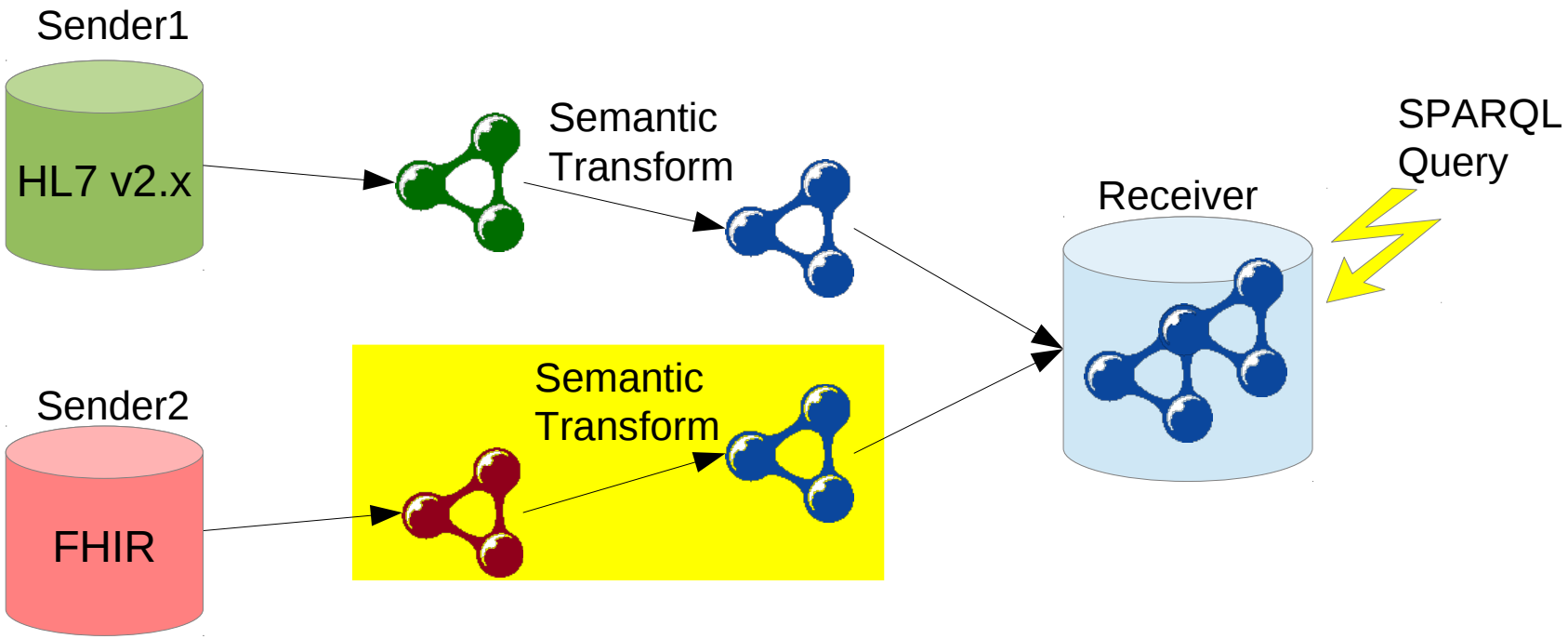
Sender1 semantic transformation

Sender1

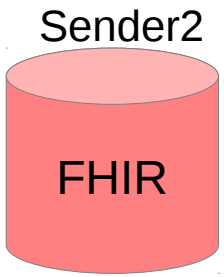

```
CONSTRUCT {
  ?observation a mOut:Observation ;
    a mOut:BP_systolic ;
    mOut:value ?value ;
    mOut:units mOut:mmHg ;
    mOut:position mOut:sitting .  }
WHERE {
  ?observation a m1:PatientObservation ;
    m1:code "3727-0" ;
    m1:value ?value ;
    m1:units "mmHg" . }
```

- RDF to RDF

Step 2: Semantic transformation



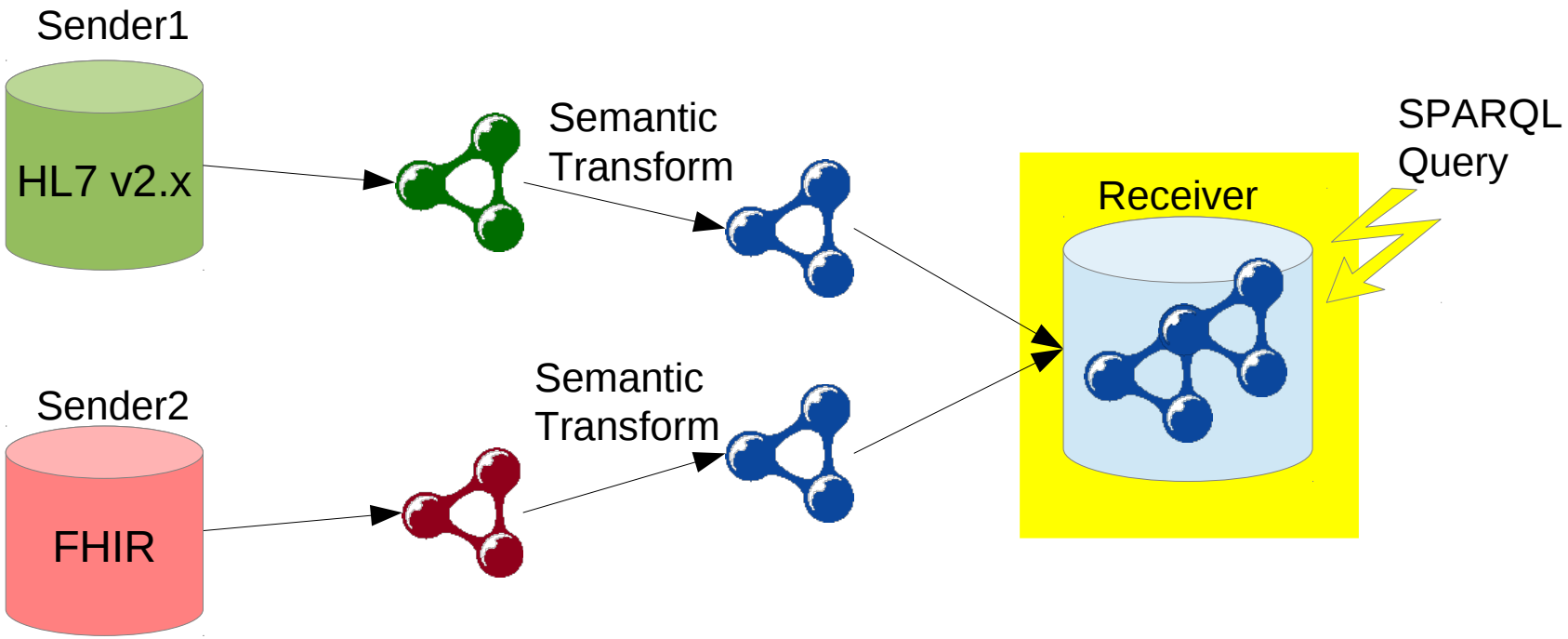
Sender2 semantic transformation




```
CONSTRUCT {
  ?observation a mOut:Observation ;
    a mOut:BP_systolic ;
    mOut:value ?value ;
    mOut:units mOut:mmHg . }
WHERE {
  ?observation a m2:Observation ;
    m2:system "http://loinc.org" ;
    m2:code "8580-6" ;
    m2:value ?value ;
    m2:units "mm[Hg]" . }
```

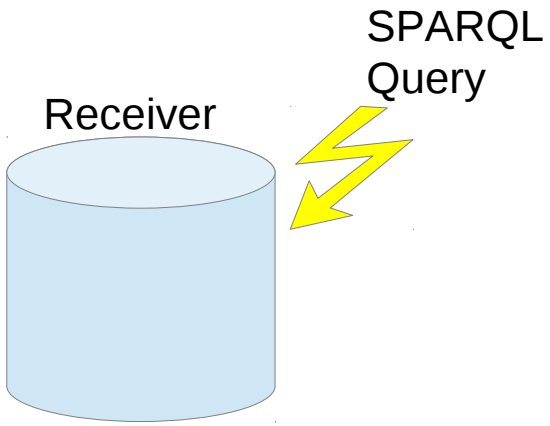
- RDF to RDF

Merged RDF


Sender1
HL7 v2.x
Semantic
Transform
SPARQL
Query
Receiver
Sender2
FHIR
Semantic
Transform

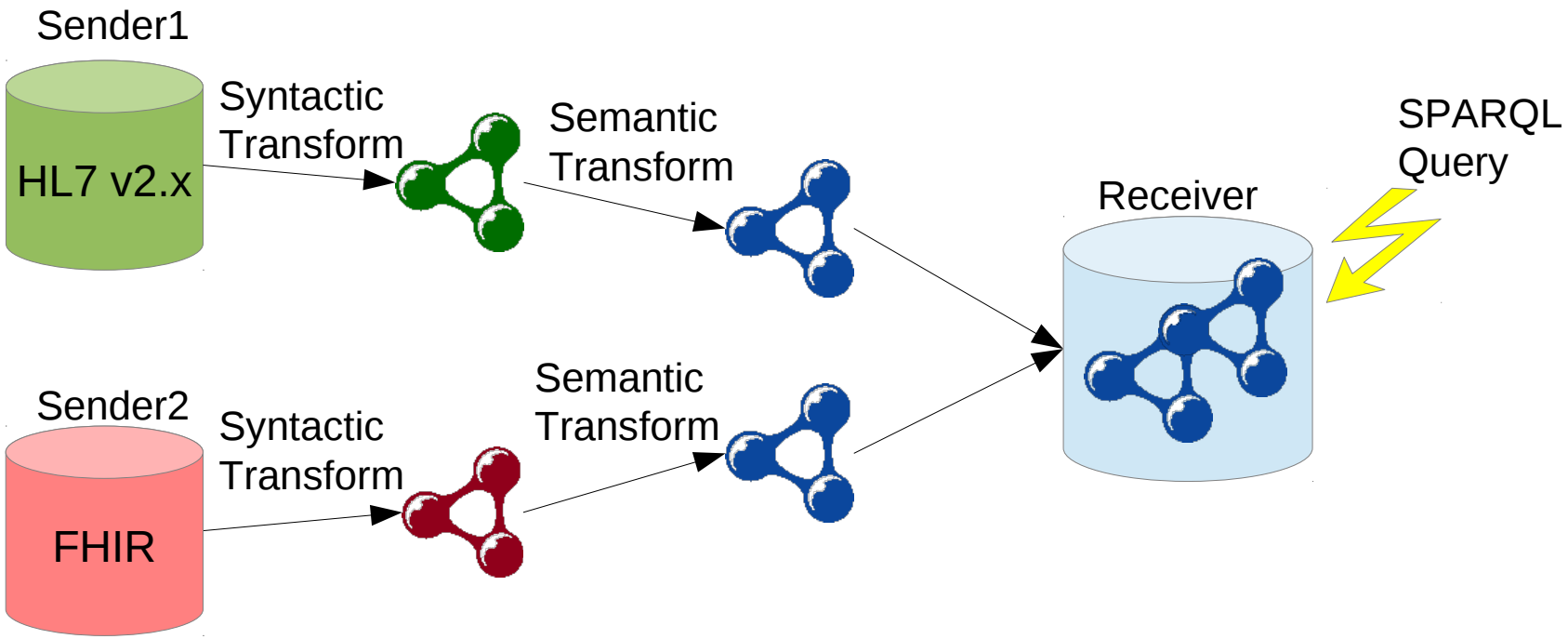
Merged RDF

```
d1:obs042 a mOut:Observation ;
  a mOut:BP_systolic ;
  mOut:value 120 ;
  mOut:units mOut:mmHg ;
  mOut:position mOut:sitting .
d2:obs-091 a mOut:Observation ;
  a mOut:BP_systolic ;
  mOut:value 107 ;
  mOut:units mOut:mmHg .
```



- Can be queried by Receiver

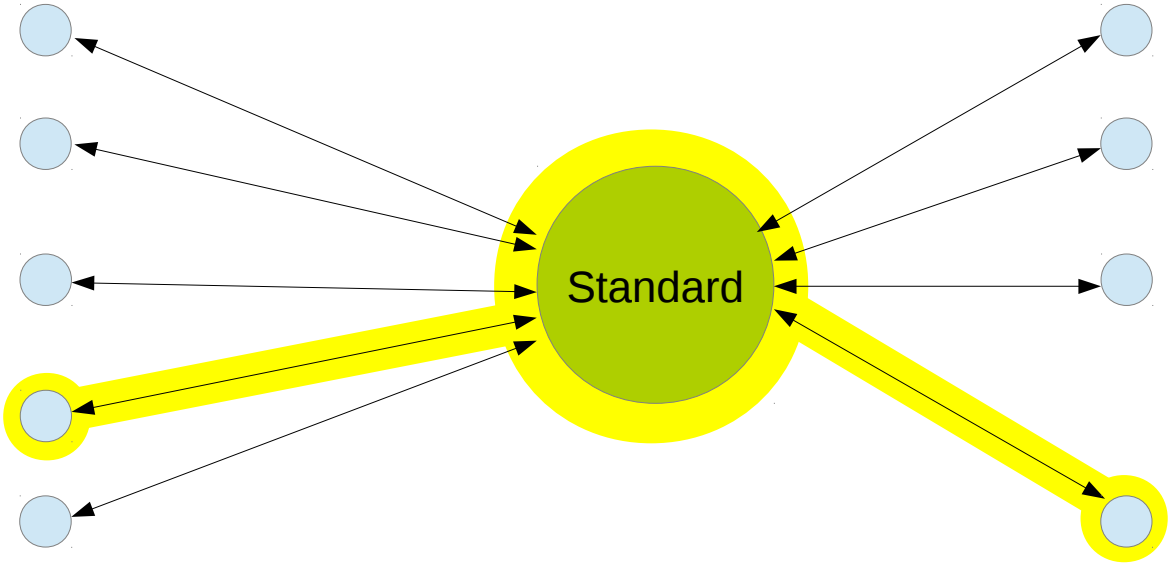
Semantic interoperability



Semantic mappings must be standardized!

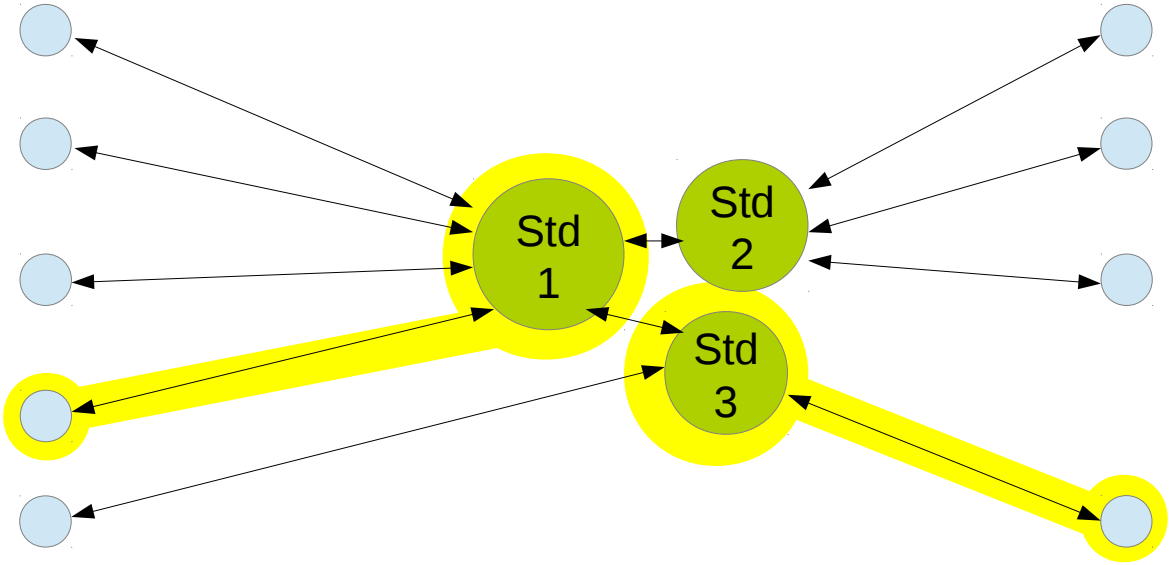
What vocabularies should be used?

Standards allow hub-and-spoke transformation



- PROS: Most efficient; desirable whenever possible
- CONS: Infeasible when committee/standard gets too big

Standards and diversity

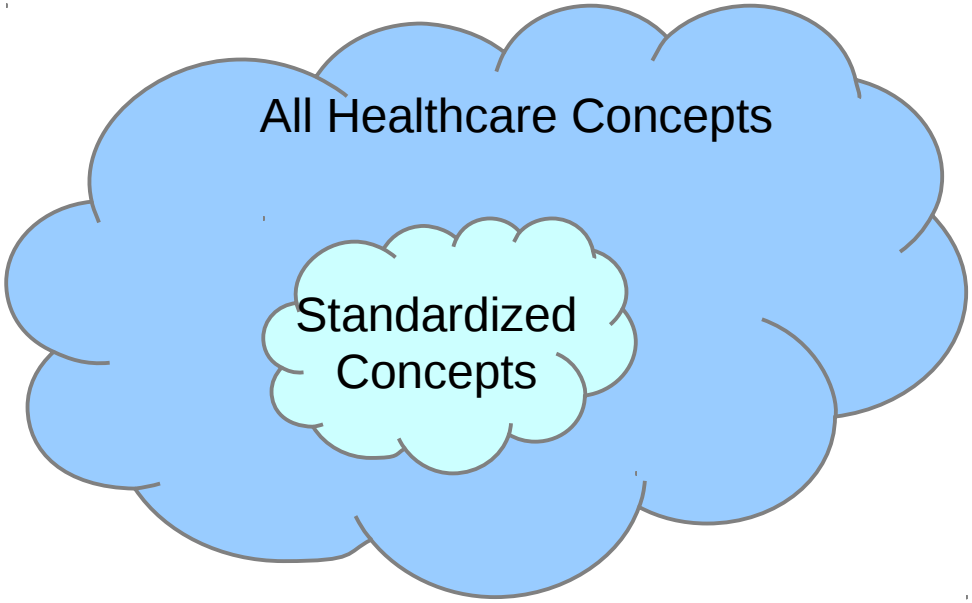


- RDF accommodates both

Evolutionary standards adoption

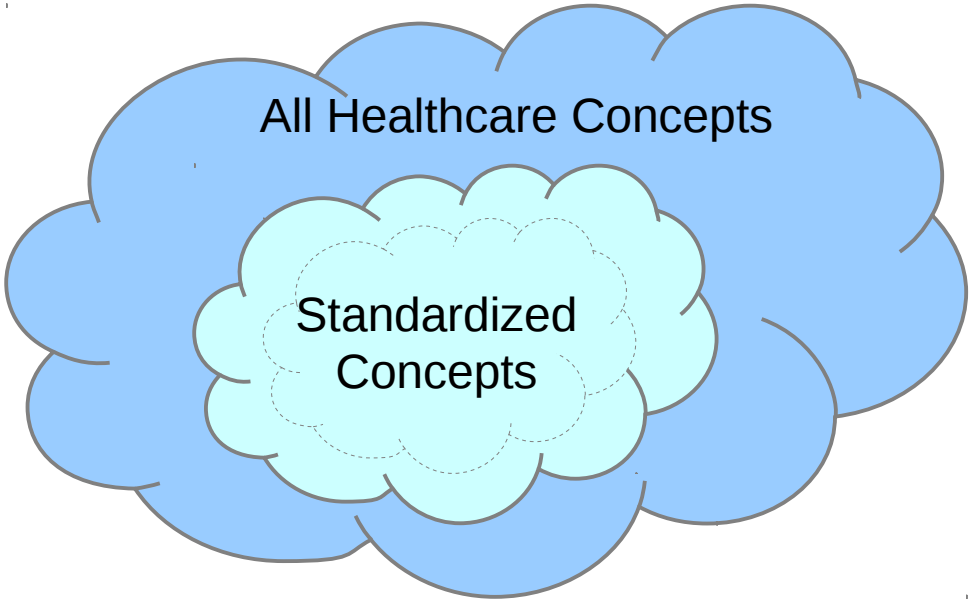
- Context:
  - Thousands of parties exchanging EHRs
  - Hundreds of different requirements
    - Clinical care, research, billing, etc.
- New standards will be created over time
- Different parties will adopt new standards at different rates

Evolutionary standards adoption



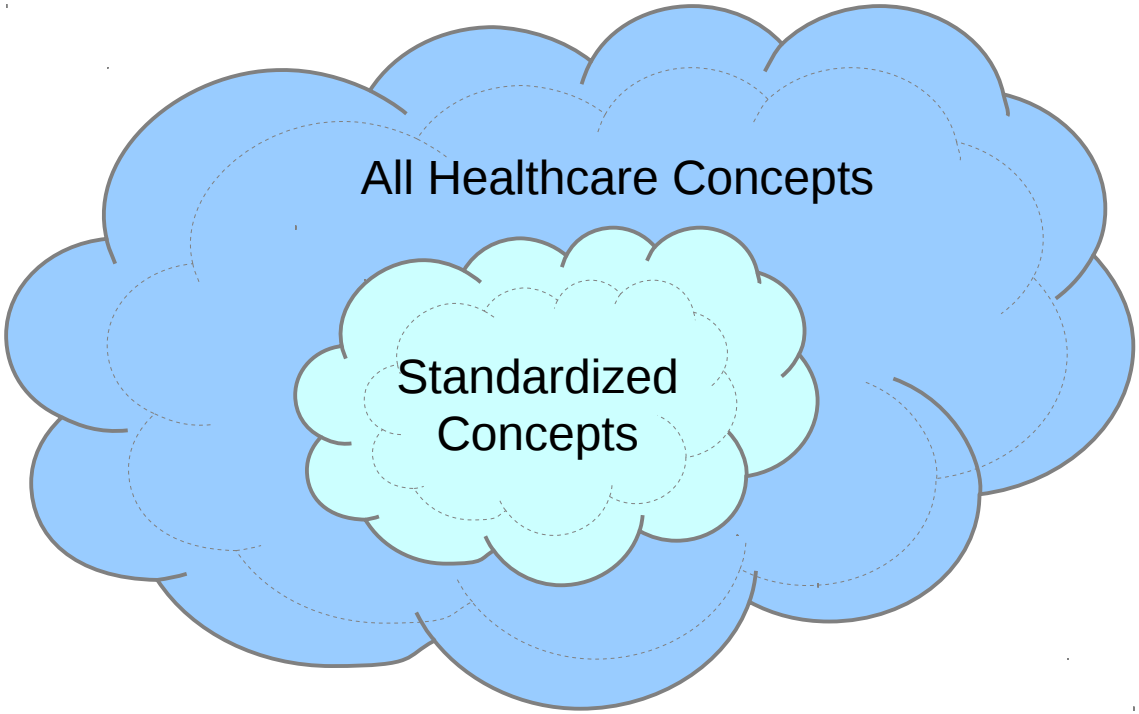
- Many concepts are not standardized

Evolutionary standards adoption



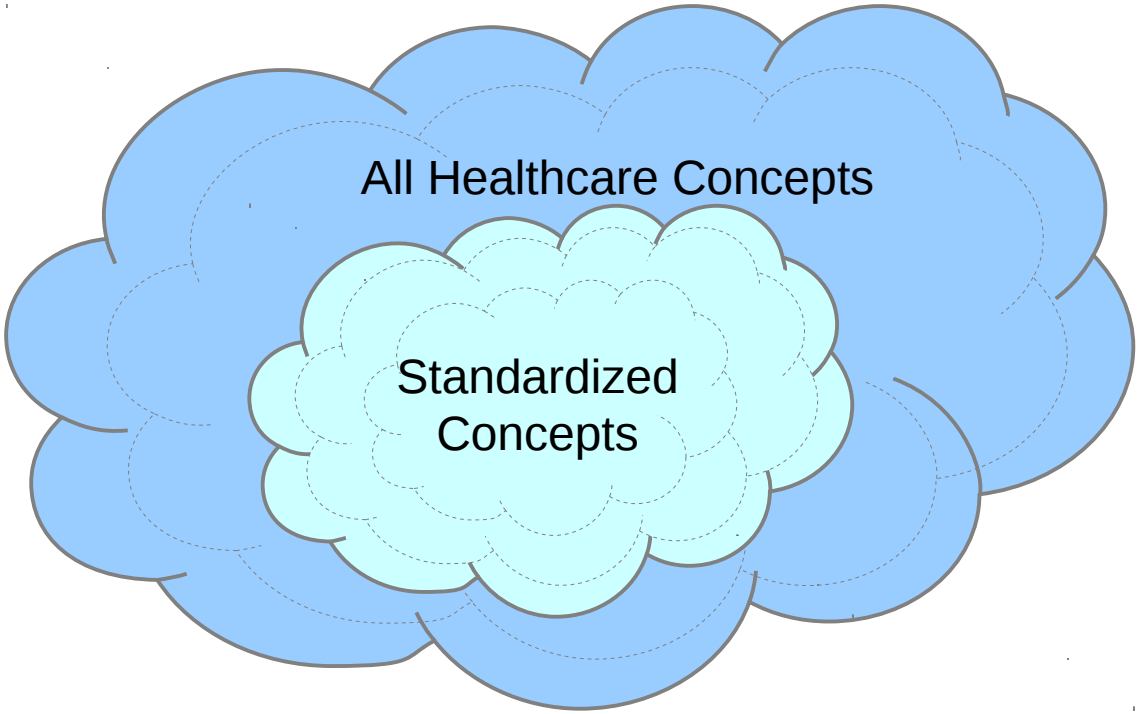
- Over time, more become standardized . . .

Evolutionary standards adoption



- But more concepts are created also

Evolutionary standards adoption



New standards must be added continually!

Semantic fidelity and granularity

- Granularity: How much detail is included?
  - E.g., "BP 120/70 mmHg, sitting, left arm" versus "BP 120/70 mmHg"
- Fidelity: Is any information lost in translation?
  - E.g., different definitions of "smoker"

Transmitted data must retain full semantic fidelity and granularity

. . . including data that is not yet standardized!

Why send non-standardized data?

- Some recipients will make use of it
  - Competitive advantage!
- Helps bootstrap standardization
  - Avoids the "no-producers-because-of-no-consumers-because-of-no-producers" dilemma

Data providers must provide all requested data!

Data must be self describing!

RDF as a Universal Healthcare Exchange Language

1. Use RDF as a *substrate* for exchange of healthcare information

2. Adopt *standard syntactic mappings* of common healthcare information formats to the RDF model

3. Adopt *standard, self-describing URIs* for healthcare concepts

4. Adopt *standard semantic mappings* between overlapping concepts

Also helpful, but beyond the scope of this workshop:
[5. Use RESTful Linked Data principles]

Questions?

KnowMED

BACKUP SLIDES

Semantic relevance is relative

- Blood Pressure measurement:
  - Sitting versus Standing
- Is the difference semantically relevant?
- Depends on the application!
- v:BP_Machine rdfs:subClassOf v:BP .
- v:BP_Manual rdfs:subClassOf v:BP .
- { ?bp a v:BP_Machine . }
  => { ?bp a v:BP . } .