Introduction to RDF

Sandro Hawke, W3C
@sandhawke

Semantic Web Tutorial
ISWC 2010

Overview

- Background
- Model
  - RDF Graphs and Triples
  - Schema
  - RDF Vocabularies
- Syntaxes
  - Turtle, RDF/XML, RDFa
  - Sparql

History

- Remember the Web in the 1990s?
  - Search was hard
  - Content labelling seemed important
- Maybe Web page metadata could help?
- Wanted to support all possible metadata
  - Page author, creator, publisher, editor, …
  - And what about them? Email? Job? Phone?
- Metadata=Data, so RDF=General Data Format

Background: URL

- We all know basic Web Addresses
  - http://google.com
  - http://www.w3.org/People/Sandro
  - https://gmail.com
- URL = Web Address of an Information Resource (Web page, image, zip file, …)

Background: URIs and IRIs

- URI = Looks the same, but might identify something else (person, place, concept)
  - Every URL is also a URI
  - Not everyone agrees with this usage
- IRI = Like URI, but not just ASCII chars
  - Every IRI can be turned into a URI (%-encoding)
  - Many of us use the term URI when we mean IRI

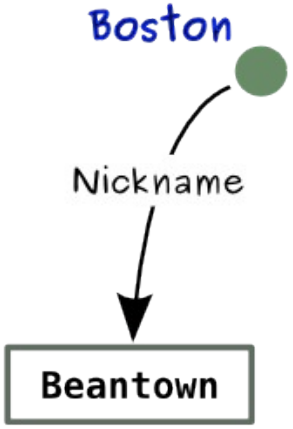
Background: QNames

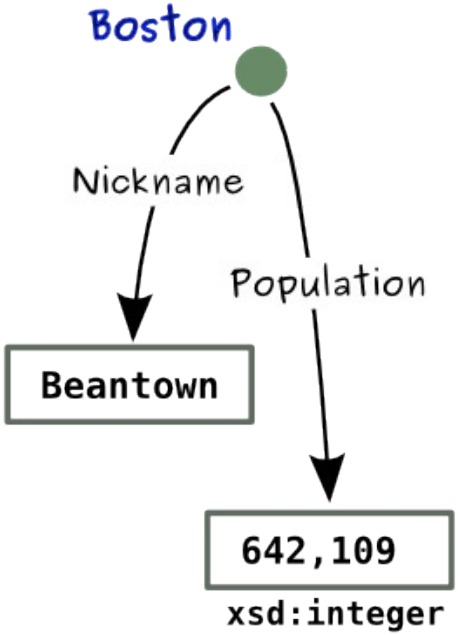
- Used in RDF as shorthand for long URIs
- If prefix “foo” is bound to http://example.com/
- Then foo:bar expands to http://example.com/bar
  - Necessary to fit any example on a page!
- Simple string concatenation
- Not quite the same as XML namespaces
- Mostly the same as CURIEs

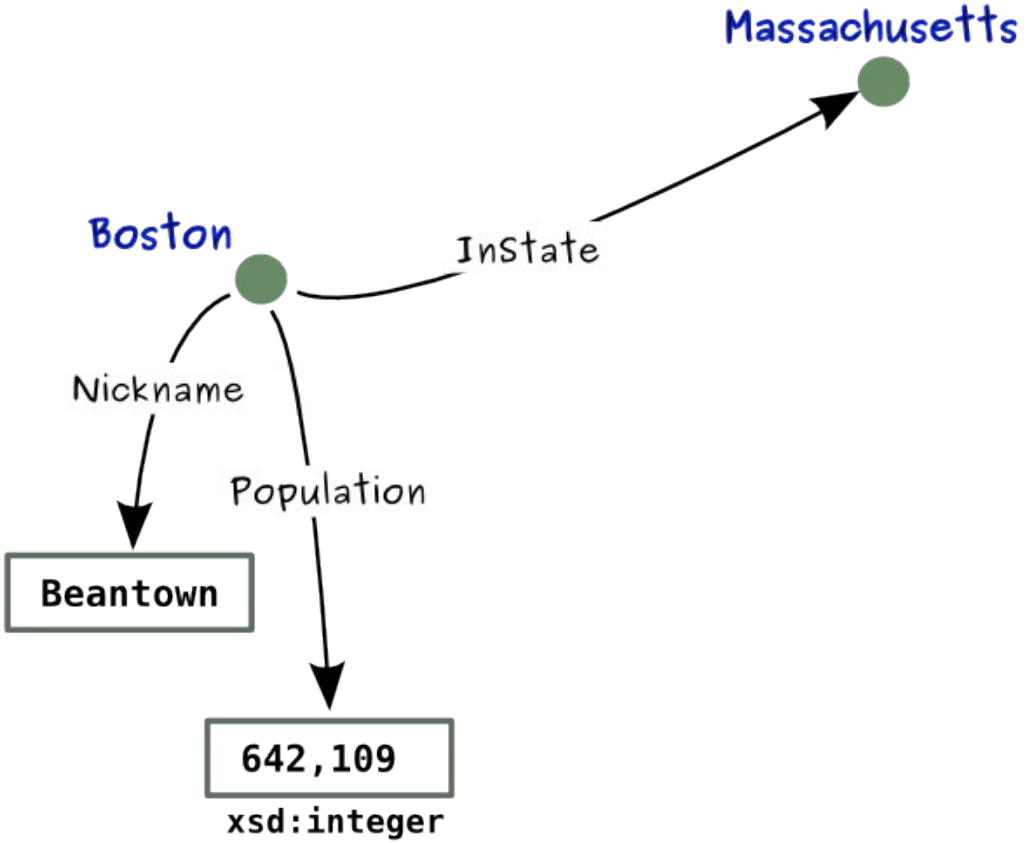
Simple, General Representation

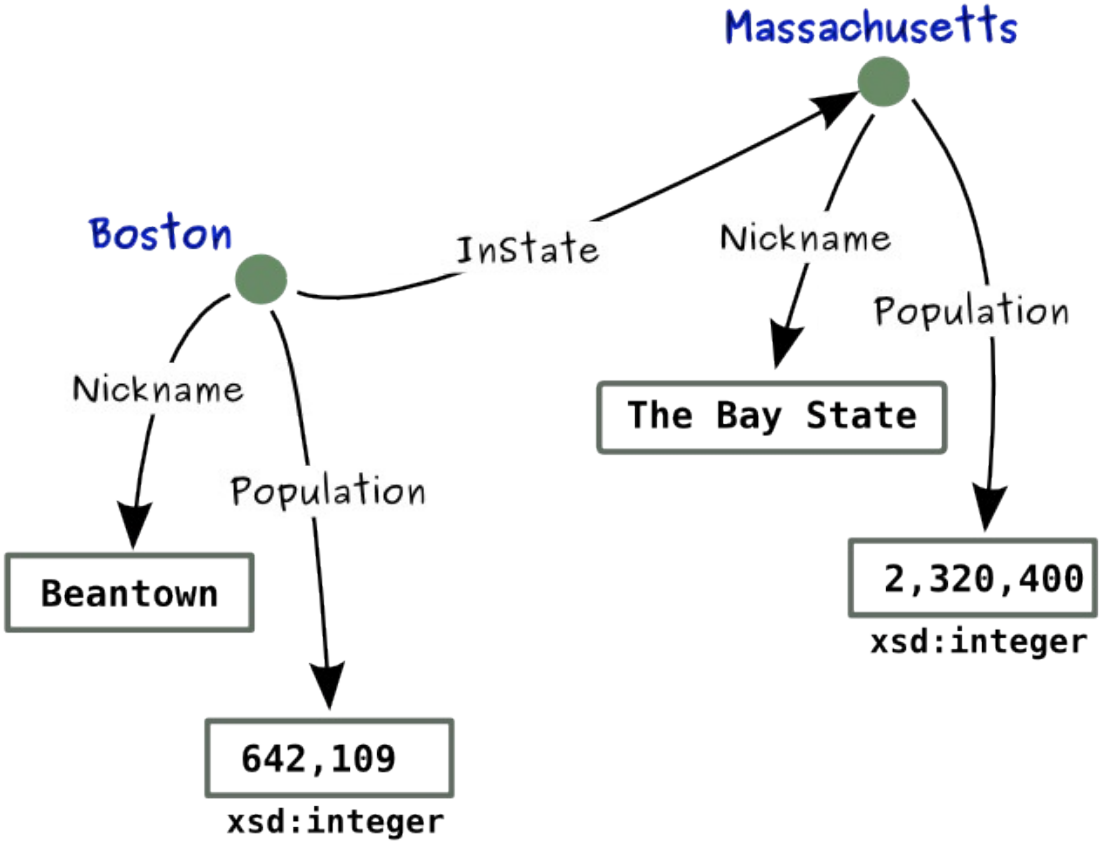
- Pick some entity as your subject
- List its attributes and values
- … and its relations to other objects

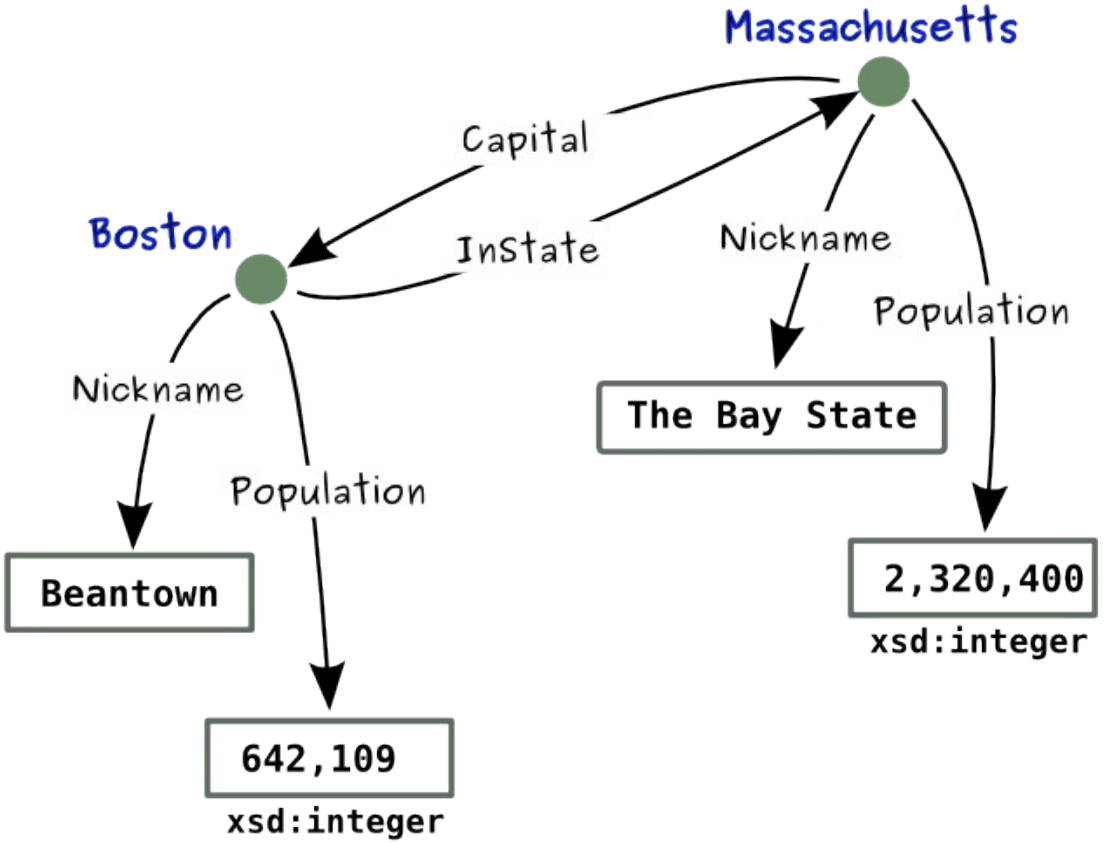
- Example subject: the City of Boston
  - Nickname: “Beantown”
  - Population: 642,109
  - In what state? Massachusetts

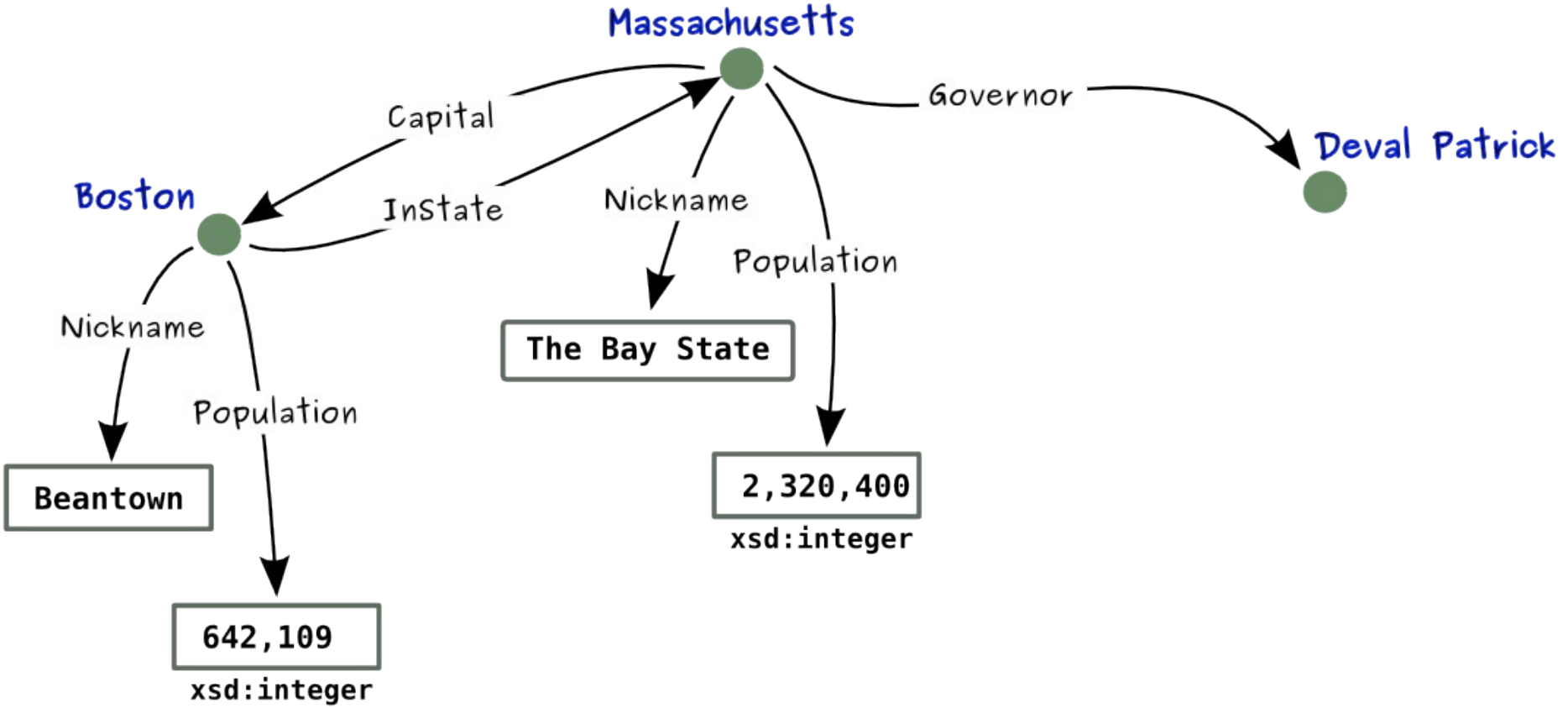

Boston
Nickname
Beantown

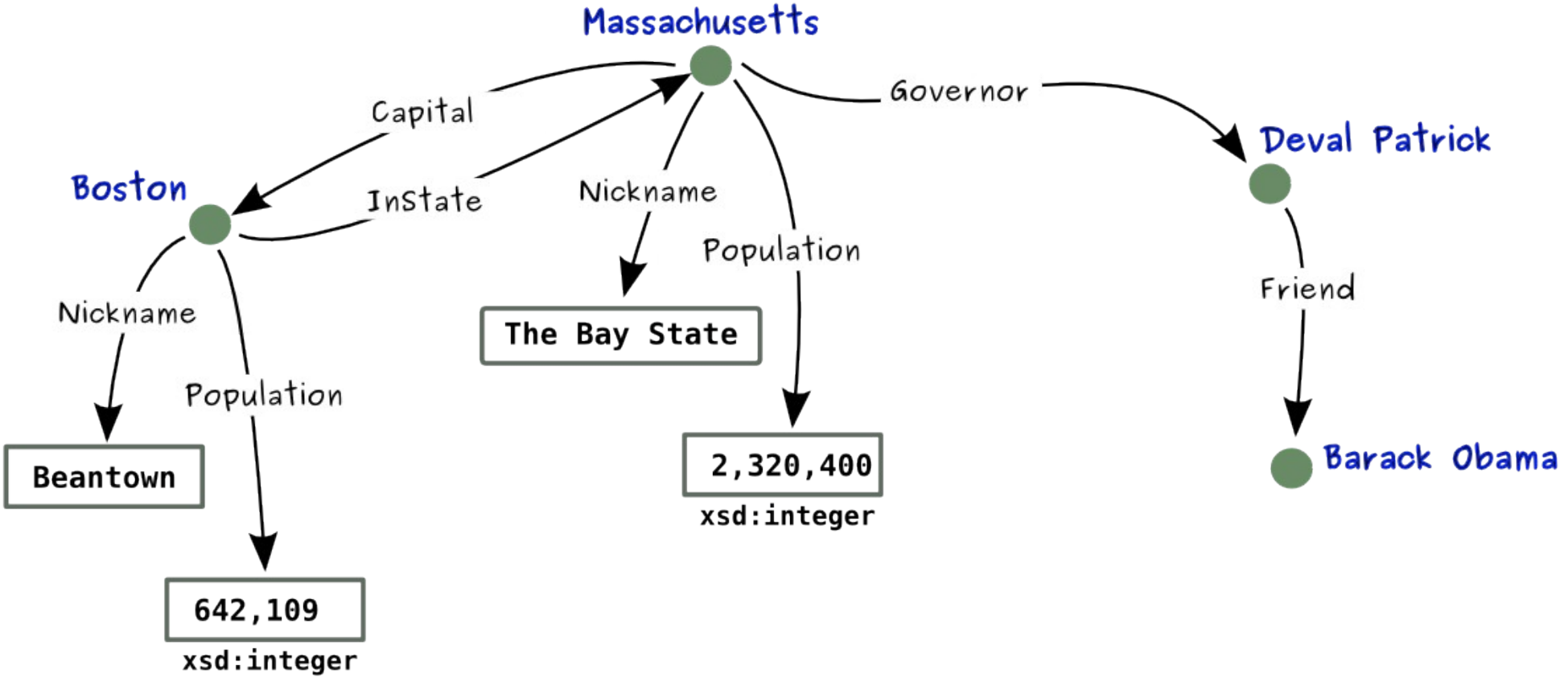

Boston
Nickname
Population
Beantown
642,109
xsd:integer

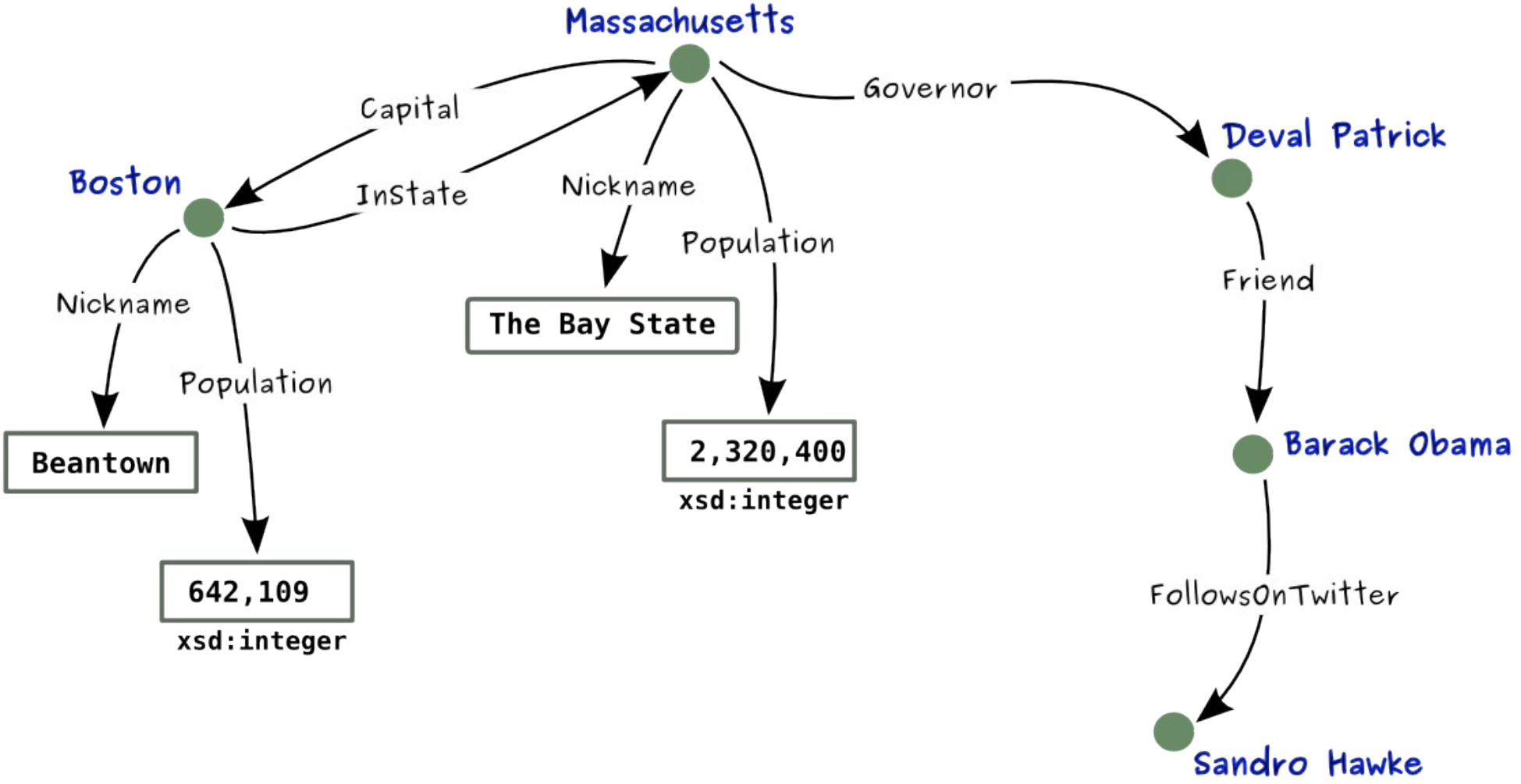

Massachusetts
Boston
InState
Nickname
Population
Beantown
642,109
xsd:integer

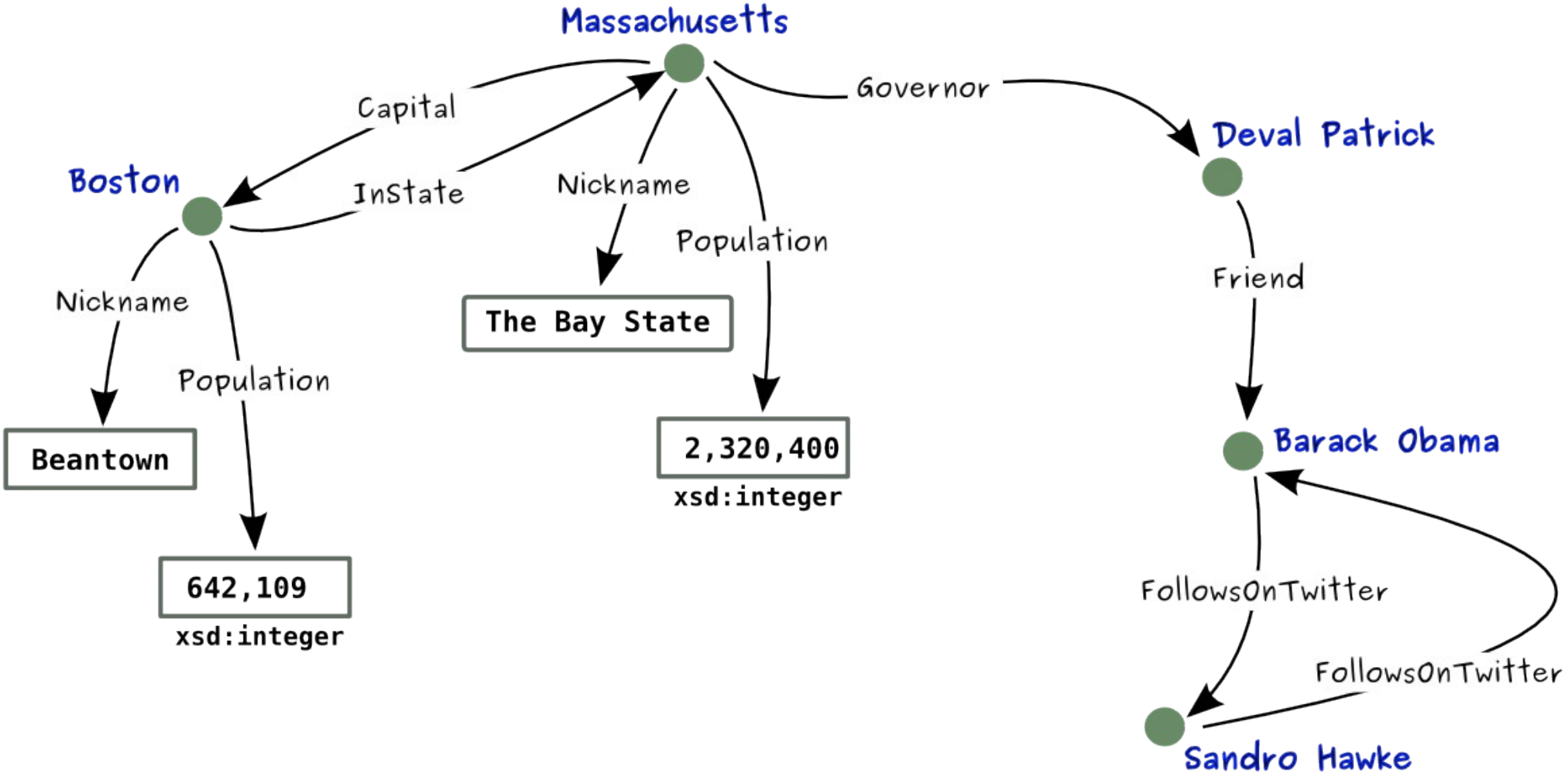

Massachusetts
Boston
InState
Nickname
Population
Nickname
Population
The Bay State
Beantown
2,320,400
xsd:integer
642,109
xsd:integer

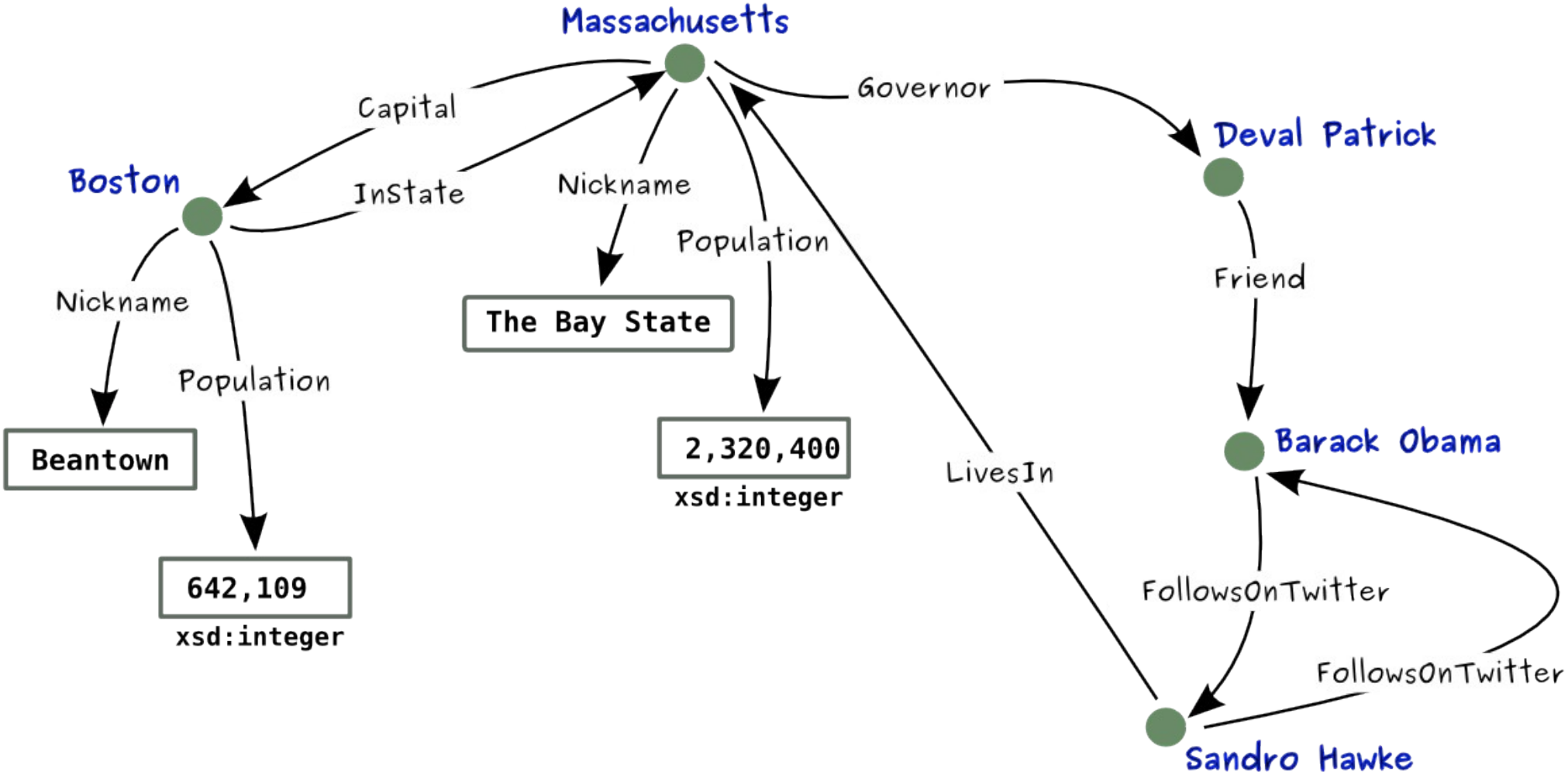

Massachusetts
Capital
Boston
InState
Nickname
Population
Nickname
The Bay State
Population
Beantown
2,320,400
xsd:integer
642,109
xsd:integer

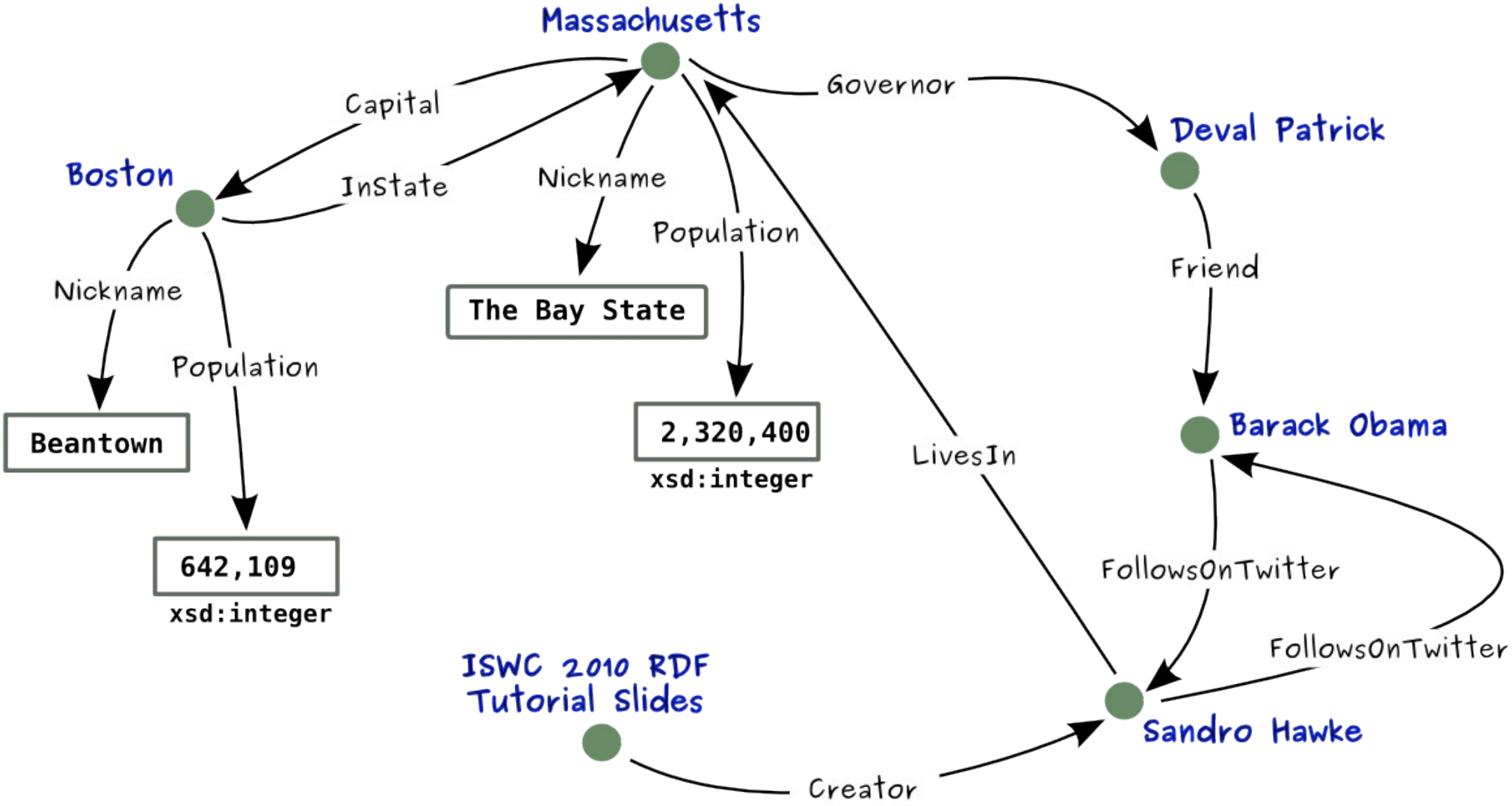

Massachusetts
Capital
Governor
Deval Patrick
Boston
InState
Nickname
Population
Nickname
The Bay State
Population
Beantown
2,320,400
xsd:integer
642,109
xsd:integer

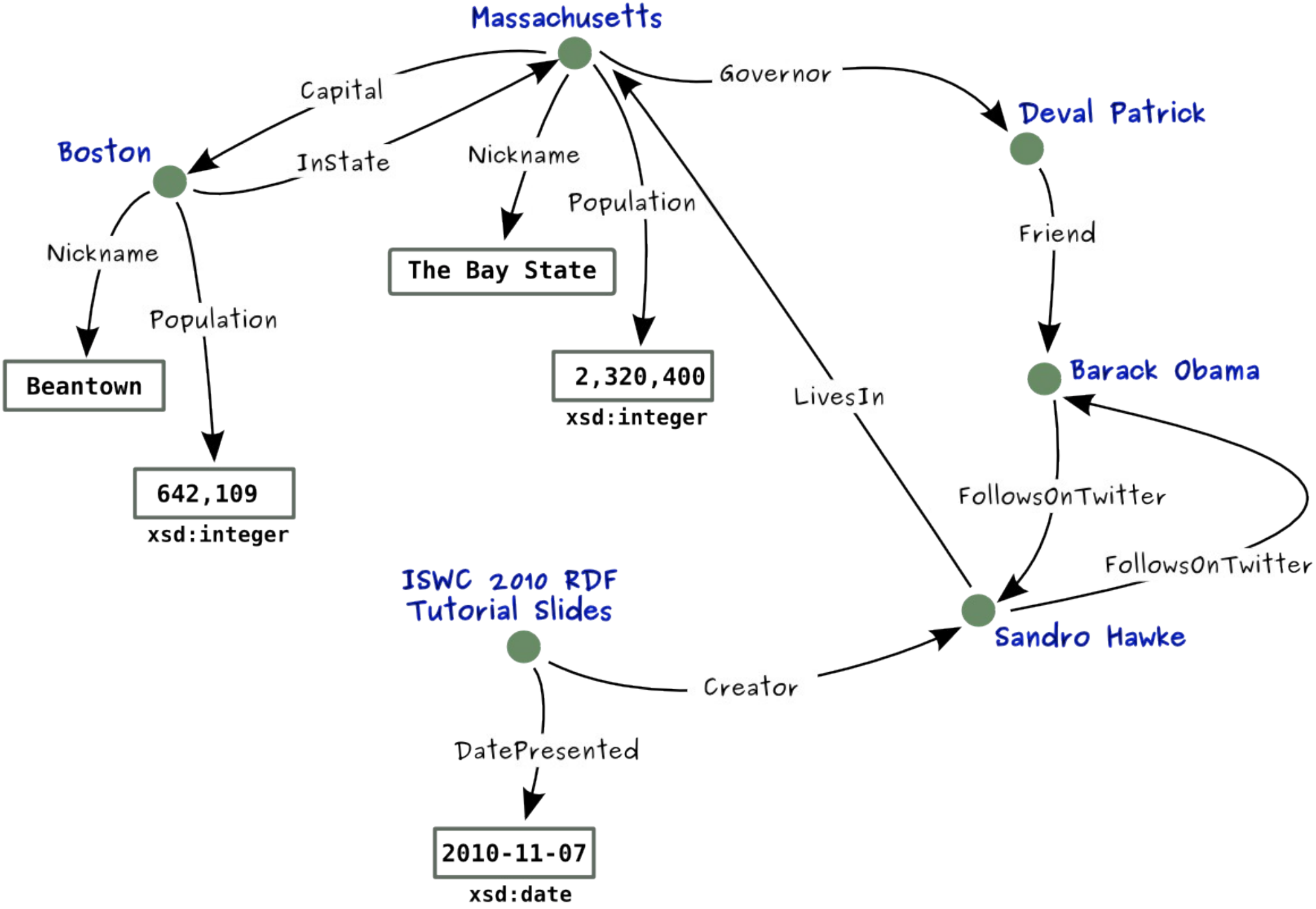

Massachusetts
Capital
Governor
Boston
InState
Nickname
Population
Deval Patrick
Friend
Nickname
The Bay State
Population
Beantown
2,320,400
xsd:integer
Barack Obama
642,109
xsd:integer


Massachusetts
Capital
Governor
Boston
InState
Nickname
Population
Deval Patrick
Nickname
The Bay State
Friend
Population
Beantown
2,320,400
xsd:integer
Barack Obama
642,109
xsd:integer
FollowsOnTwitter
Sandro Hawke


Massachusetts
Capital
Governor
Boston
InState
Nickname
Population
Deval Patrick
Nickname
The Bay State
Friend
Population
Beantown
2,320,400
xsd:integer
Barack Obama
642,109
xsd:integer
FollowsOnTwitter
FollowsOnTwitter
Sandro Hawke


Massachusetts
Capital
Boston
InState
Nickname
Population
Governor
Deval Patrick
Nickname
Population
The Bay State
Friend
Beantown
2,320,400
xsd:integer
Barack Obama
LivesIn
642,109
xsd:integer
FollowsOnTwitter
FollowsOnTwitter
Sandro Hawke


Massachusetts
Capital
Governor
Boston
InState
Nickname
Population
Deval Patrick
Friend
Nickname
The Bay State
Population
Beantown
2,320,400
xsd:integer
Barack Obama
LivesIn
Population
642,109
xsd:integer
FollowsOnTwitter
FollowsOnTwitter
ISWC 2010 RDF
Tutorial Slides
Sandro Hawke
Creator

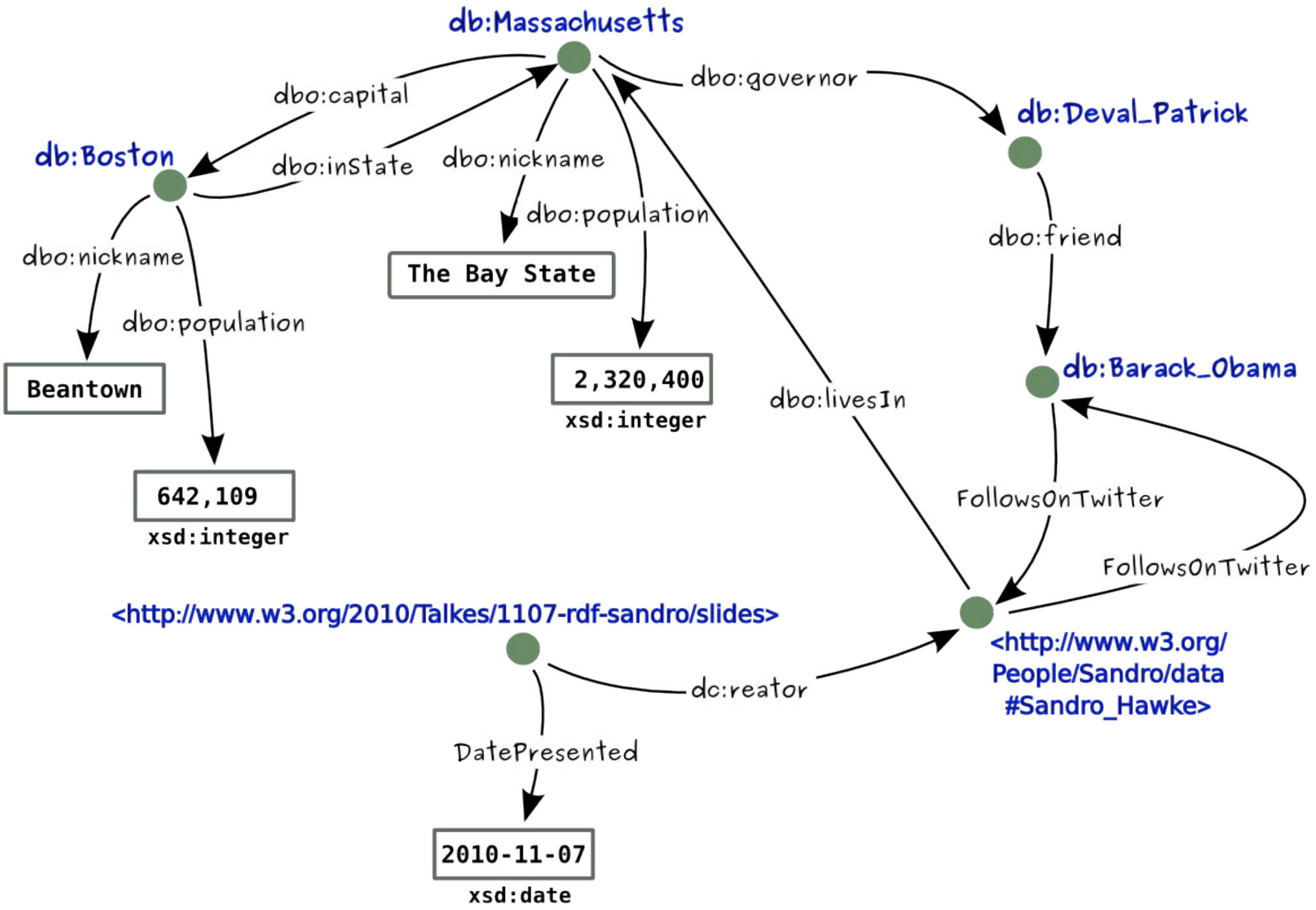

Massachusetts
Capital
InState
Nickname
Population
Governor
Boston
Deval Patrick
Nickname
Population
Friend
The Bay State
Beantown
2,320,400
xsd:integer
642,109
xsd:integer
LivesIn
Barack Obama
FollowsOnTwitter
FollowsOnTwitter
ISWC 2010 RDF
Tutorial Slides
Sandro Hawke
Creator
DatePresented
2010-11-07
xsd:date

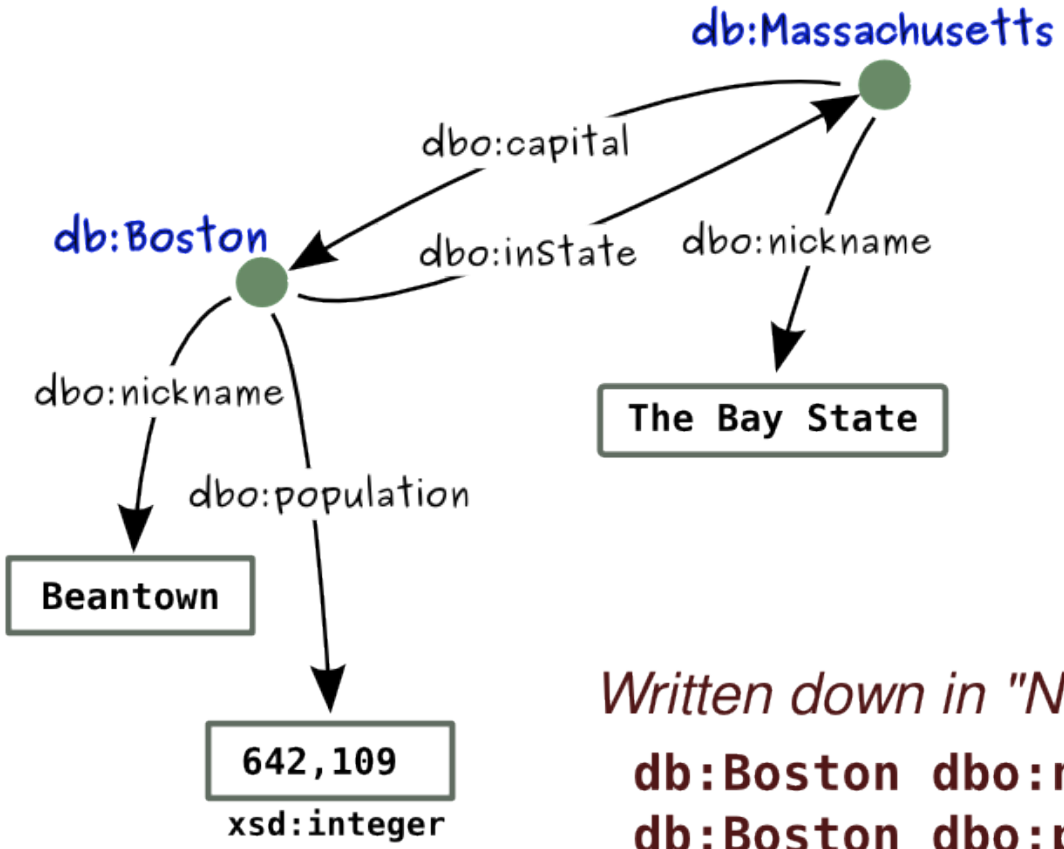
Unambiguous Names

- How many things are named “Boston”?
- How about “Riverside”?

- So, we use URIs. Instead of “Boston”:
  - http://dbpedia.org/resource/Boston
  - QName: db:Boston

- And instead of “nickname” we use:
  - http://example.org/terms/nickname
  - QName: dbo:nickname


db:Massachusetts
dbo:capital
dbo:governor
db:Boston
dbo:inState
dbo:nickname
db:Deval_Patrick
dbo:population
dbo:friend
dbo:nickname
The Bay State
dbo:population
Beantown
2,320,400
xsd:integer
dbo:livesIn
db:Barack_Obama
642,109
xsd:integer
FollowsOnTwitter
FollowsOnTwitter
<http://www.w3.org/2010/Talkes/1107-rdf-sandro/slides>
<http://www.w3.org/People/Sandro/data#Sandro_Hawke>
dc:reator
DatePresented
2010-11-07
xsd:date



Written down in "N-Triples":

```
db:Boston dbo:nickname "Beantown".
db:Boston dbo:population "642109"^^xsd:integer.
db:Boston dbo:inState db:Massachusetts.
db:Massachusetts dbo:capital db:Boston.
db:Massachusetts dbo:nickname "The Bay State".
```

Subject **Predicate (Property)** **Object (Value)**

RDF “Literals”

- Data values
- Often shown inside a rectangle in graph pictures
- Plain Literals
  - Just strings, “Hello, World”
- Language-Tagged Literals
  - “Bonjour, Monde”@fr
- XML Schema Types
  - “3.14”^^xs:float

Nodes with URI Labels

- If the thing represented by the node has a URI, use it as a label for the node.
  - We often just write qnames
  - Put URIs in <brackets> to distinguish them

- <http://www.w3.org> ns:created “1994-04-15”^^xsd:date.
- <http://www.w3.org/People/Sandro/data#Sandro_Hawke> foaf:firstName “Sandro”.

Blank Nodes

- Nodes with no URI, also called “bnodes”
- For when you don't have a URI for something
- … and don't want to create one

- In N-Triples:

ns1:sandro foaf:knows _:node1.

_:node1 foaf:name “Dan Brickley”.

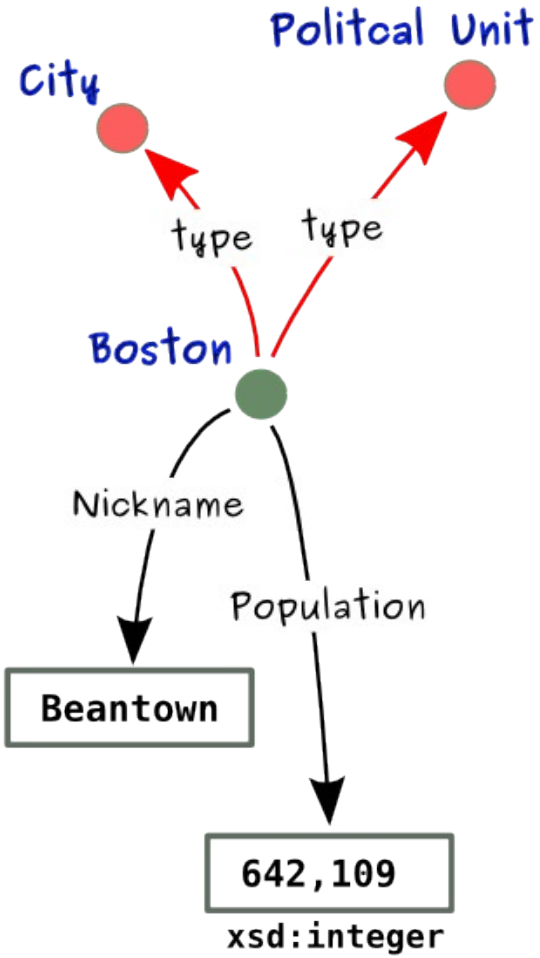
_:node1 foaf:mbox <mailto:danbri@danbri.org>.

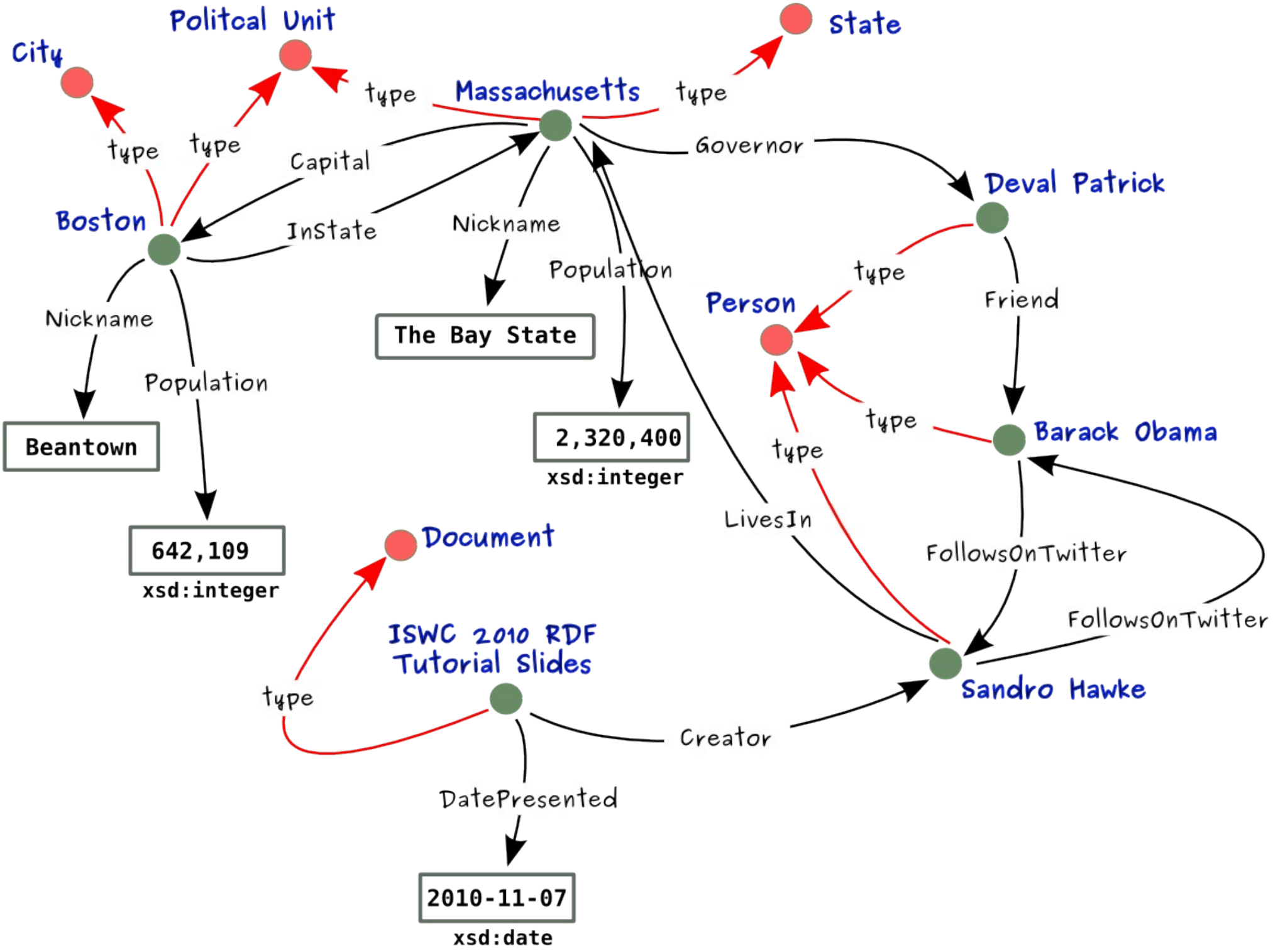
Properties

- The “Predicate” or “Property”
  - Attribute, Relation
  - Always named with a URI
- Same URI can be used as Subject or Object
  - This allows self-description, documentation

Classes and rdf:type

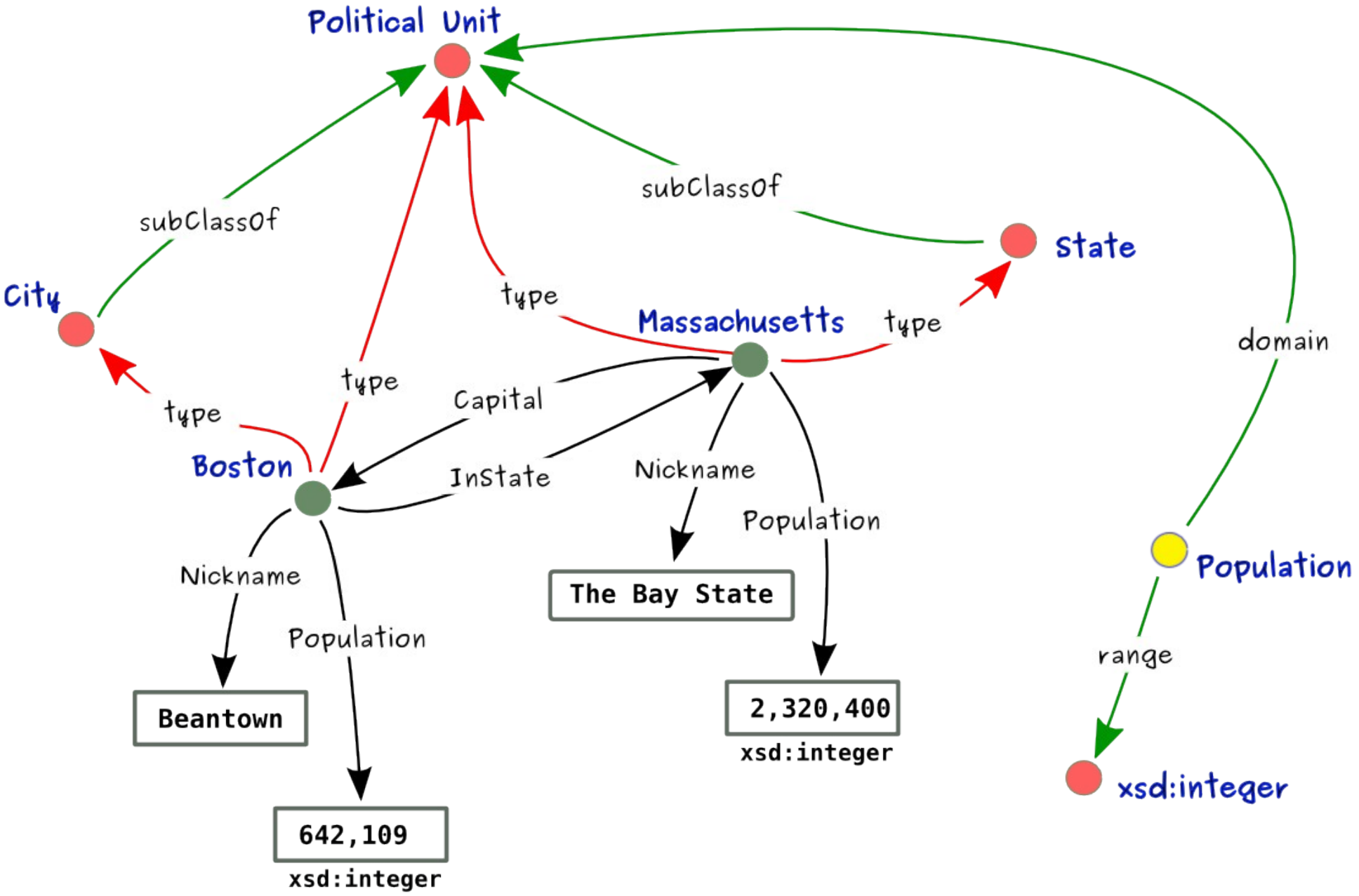
- Sometimes it's helpful to organize using types
- We can attach types using “type” arcs
- … and then use those in lots of ways, later


Politcal Unit
City
type
type
Boston
Nickname
Population
Beantown
642,109
xsd:integer


Politcal Unit
State
City
type
Massachusetts
type
type
type
Capital
Governor
Boston
InState
Nickname
Deval Patrick
Population
type
Person
Friend
Nickname
The Bay State
Population
type
Beantown
2,320,400
xsd:integer
Barack Obama
642,109
xsd:integer
Document
LivesIn
FollowsOnTwitter
FollowsOnTwitter
ISWC 2010 RDF
Tutorial Slides
type
Sandro Hawke
Creator
DatePresented
2010-11-07
xsd:date

A Little RDF Schema

- X rdfs:subclassOf Y
  - Everything of type X is also of type Y
  - Dog rdfs:subclassOf Animal, Spot rdf:type Dog |= Spot rdf:type Animal
- X rdfs:domain Y
  - Everything that has an X property is of type Y
  - ownsPet rdfs:domain Human, Sam ownsPet Spot |= Sam rdf:type Human
- X rdfs:range Y
  - Every value of an X property is of type Y
  - OwnsPet rdfs:range Animal, Sam ownsPet Spot |= Spot rdf:type Animal


Political Unit
subClassOf
subClassOf
City
State
type
type
Massachusetts
type
domain
type
type
Capital
Boston
InState
Nickname
Population
Nickname
Population
The Bay State
Population
range
Beantown
2,320,400
xsd:integer
xsd:integer
642,109
xsd:integer

Vocabularies

- Often formalized with Schemas or **Ontologies**
- RDF, RDF Schema
  - rdf:type, rdfs:subClassOf, rdfs:comment
- Friend of a Friend
  - foaf:name
- Dublin Core
  - dc:creator, dcterms:temporal
- Good Relations
  - gr:ProduceOrServiceModel, ...

Turtle

- Very simple RDF Syntax
  - N-Triple plus a few bits of syntax sugar
- De facto standard now
  - Widely implement
  - Should be W3C Recommendation soonish

```
db:Boston dbo:nickname “Beantown”;
          dbo:population “610000”^^xs:integer;
          dbo:inState db:Massachusetts.
db:Massachusetts ...
```

RDF/XML

- W3C Standard since 1999, revised in 2004
- Used to be the only standard
- Can look like “normal” XML, but works differently

RDF/XML

```
<rdf:RDF>
   <Description rdf:about="http://dbpedia.org/resource/Boston">
      <nickname>Beantown</nickname>




   </Description>
</rdf:RDF>
```

RDF/XML

```
<rdf:RDF>
   <Description rdf:about=”http://dbpedia.org/resource/Boston”>
      <nickname>Beantown</nickname>
       <population
            rdf:datatype=”xs:integer”>610104</dbo:population>




    </Description>
</rdf:RDF>
```

RDF/XML

```
<rdf:RDF>
   <Description rdf:about="http://dbpedia.org/resource/Boston">
      <nickname>Beantown</nickname>
      <population
            rdf:datatype="xs:integer">610104</dbo:population>
      <inState>
       <Description rdf:about="http://dbpedia.org/resource/Massachusetts">
          <nickname>The Bay State</nickname>
          ....
       </Description>
      </inState>
    </Description>
</rdf:RDF>
```

RDFa

- **RDF** triples in XHTML **a**ttributes
- W3C Recommendation 2008
- RDFa 1.1 underway
- Build easily on existing HTML pipeline
- In some case, just means adding a few attributes

RDFa Example

```
<div about="http://dbpedia.org/resource/Boston"
      xmlns:dbo="http://example.com/dbo/">
Boston has the nickname
  <span property="dbo:nickname">Beantown</span>
```

RDFa Example

```
<div about="http://dbpedia.org/resource/Boston"
      xmlns:dbo="http://example.com/dbo/">
Boston has the nickname
  <span property="dbo:nickname">Beantown</span>
and a population of
  <span property="dbo:population
        datatype="xs:integer">642109</span>.
</div>
```

RDFa Example

```
<div about="http://dbpedia.org/resource/Boston"
      xmlns:dbo="http://example.com/dbo/">
Boston has the nickname
  <span property="dbo:nickname">Beantown</span>
and a population of
  <span property="dbo:population
          datatype="xs:integer">642109</span>.
It is located in
  <a rel="dbo:inState"
     href="http://dbpedia.org/resource/Massachusetts">Massachusetts</a>
</div>
```

RDFa Example

<div about="http://dbpedia.org/resource/Boston"
xmlns:dbo="http://example.com/dbo/">
Boston has the nickname
Beantown
and a population of
<span property="dbo:population
datatype="xs:integer">642109.
It is located in
<a rel="dbo:inState"
href="http://dbpedia.org/resource/Massachusetts">Massachusetts
Which has
** **
the nickname
The Bay State
....
** **

SPARQL

- Language for querying collection of RDF Graphs
- Somewhat like SQL
- W3C Recommendation in 2008
- V1.1 will add update, be more expressive

```
PREFIX foaf:   <http://xmlns.com/foaf/0.1/>
SELECT ?name ?mbox
WHERE
  { ?x foaf:name ?name .
    ?x foaf:mbox ?mbox }
```

Summary of Model

- RDF started as metadata
- It's a general data format, a simple KR
- A collection of RDF knowledge is
  - A graph of subject/object nodes and property arcs
  - Nodes may be labeled with URIs, or Blank
  - Leaf nodes may be literals, optionally typed
- Vocabularies (Ontologies)
  - Classes, Properties, Individuals
  - Each with a well-known URI

Summary of Syntaxes

- An RDF Graph can be serialized many ways
  - Turtle (N-Triples, N3) very simple, a de facto standard
  - RDF/XML is the original standard. It's XML, but has some impedance mismatch with XML tools
  - RDFa is good for RDF in HTML
  - Other syntaxes exist, might be standardized
    - Eg JSON
  - RDF can also be accessed via APIs and SPARQL

More Information

- Me:
  - Sandro Hawke, sandro@w3.org
  - @sandhawke on twitter
- Semantic Web / RDF
  - http://www.w3.org/standards/semanticweb/
  - http://www.w3.org/RDF/
- This Talk
  - http://www.w3.org/2010/Talks/1107-rdf-sandro