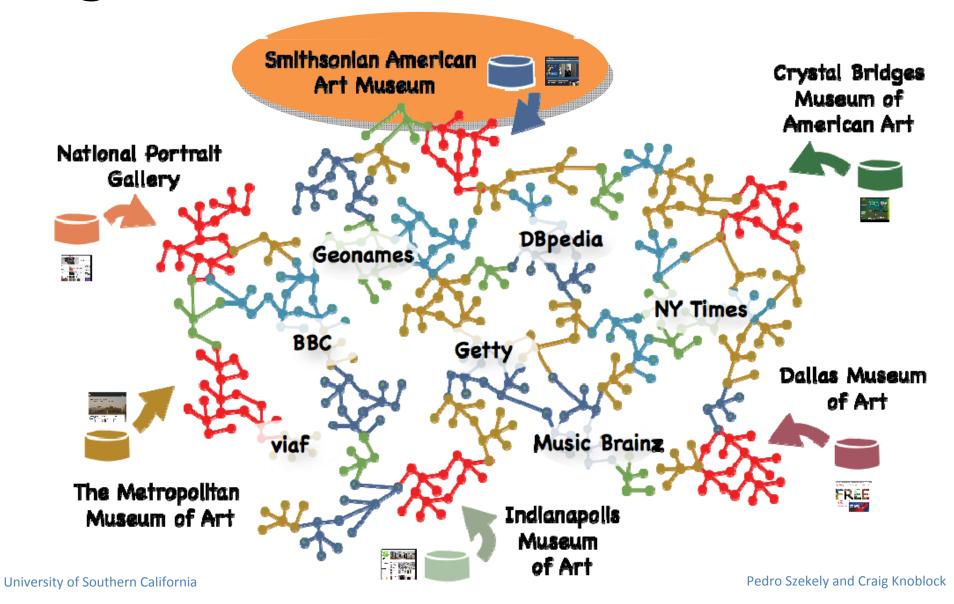
# Connecting the Smithsonian American Art Museum to the Linked Data Cloud

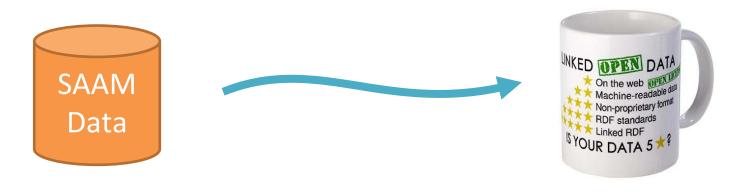
Pedro Szekely, Craig A. Knoblock, Fengyu Yang, Xuming Zhu, Eleanor E. Fink, Rachel Allen, and Georgina Goodlander

University of Southern California, Los Angeles, California, USA Nanchang Hangkong University, Nanchang, China Smithsonian American Art Museum, Washington, DC, USA

## Big Picture



### **Problem**



What ontology to use?

Data consistency

What to link to?

**Structure** mismatches

100% precision

#### How to enable museums to do this themselves?

### Steps to Create Linked Data

- Map data to RDF
  - ... select ontologies
  - ... define mappings
- Link to external resources
  - ... identify the links
- Curate the Linked Data
  - ... museums demand 100% correctness

## select ontologies



Definition of the **CIDOC Conceptual Reference Model** 

Produced by the ICOM/CIDOC Documentation Standards Group, Continued by the CIDOC CRM Special Interest Group

Version 5.0.4

November 2011



Europeana Data Model Mapping Guidelines v1.0.1



Europeana Data Model Primer



61 pages

Definition of the Europeana Data Model elements

Version 5.2.3, 24/02/2012

Europeana v1.0

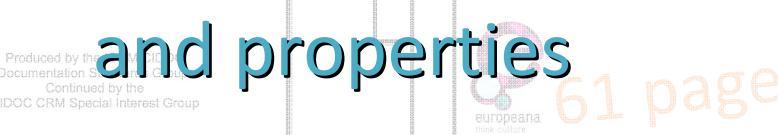




## Complicated

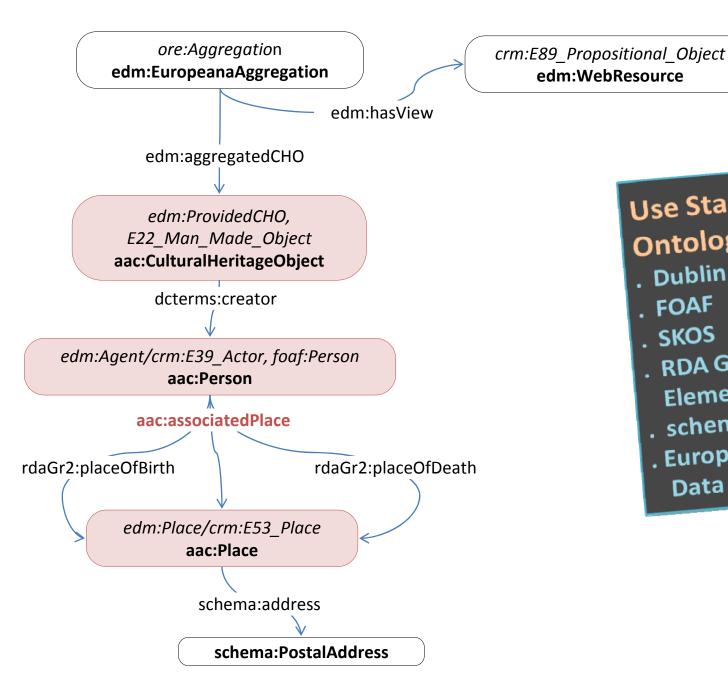
Many irrelevant classes

November 2011

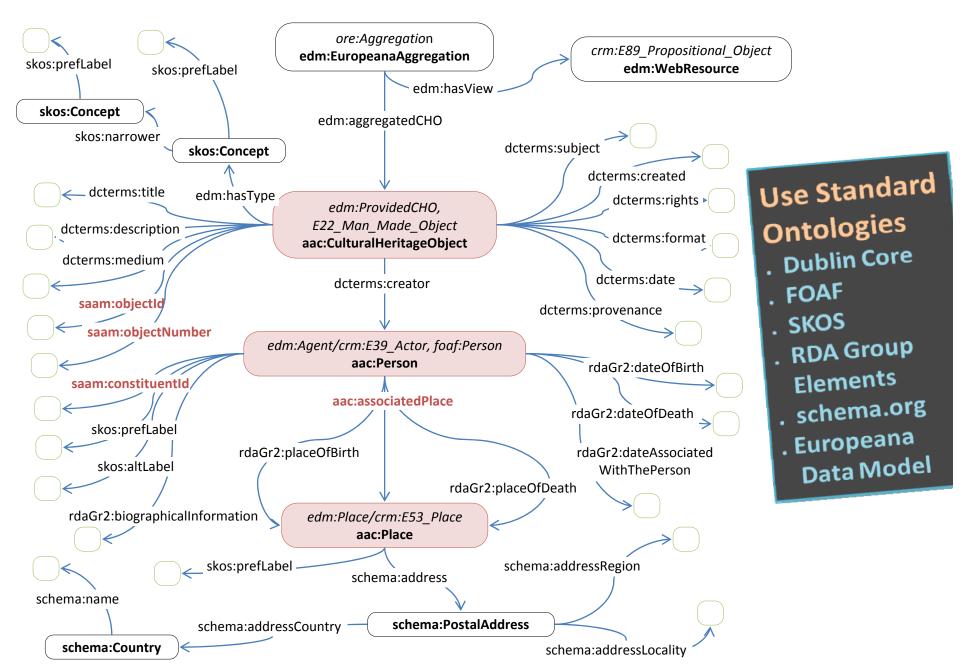


Definition of the Europeana Data Model elements

Europeana v1.0



**Use Standard** Ontologies . Dublin Core . FOAF . SKOS . RDA Group Elements . schema.org . Europeana Data Model

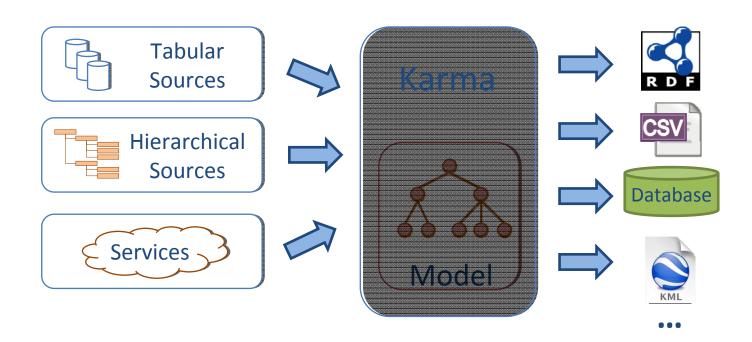


## mapping the data to the ontologies

how to enable museums to do this themselves?

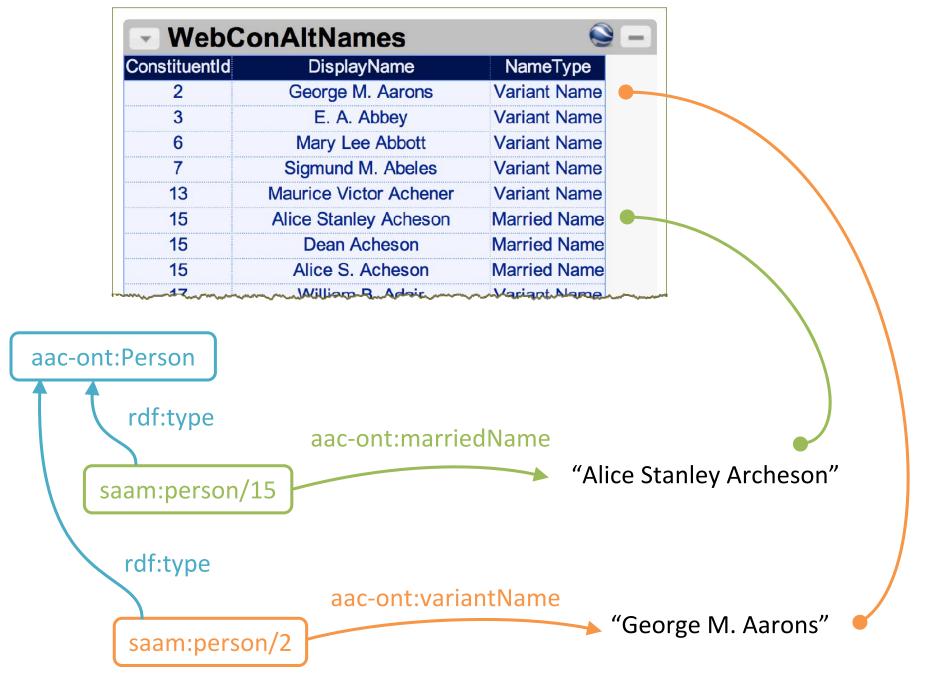
#### Karma

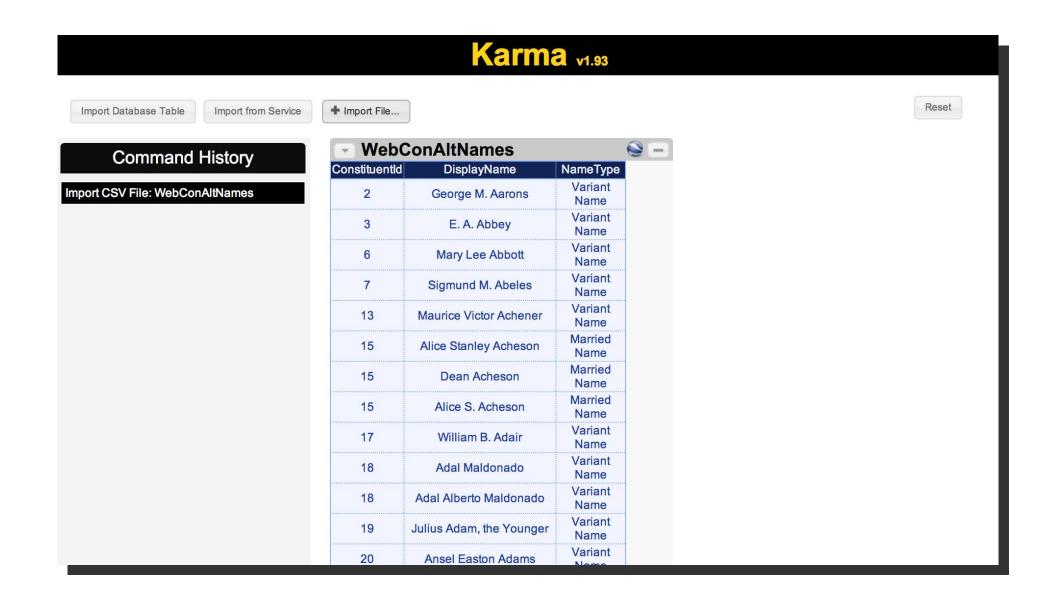
Interactive tool for rapidly extracting, cleaning, transforming, integrating, and publishing data



[ Knoblock, Szekely, et al. Semi-automatically mapping structured sources into the semantic web. ISWC 2012 ]

# specifying transformations and mapping to properties with Karma





http://isi.edu/integration/karma

# mapping to object properties using Karma



http://isi.edu/integration/karma

University of Southern California Pedro Szekely and Craig Knoblock

## **Evaluation of Data Mapping Using Karma**

SAAM database

8 tables

29 columns

**Ontologies** 

407 classes

105 data properties

229 object properties

	# of times Karma's top 4 suggestions contain the correct semantic type	# of times Karma correctly assigns object properties	Time (minutes)
Run 1: no training data	7 out of 29 (24%)	30 out of 35 (85%)	18
Run 2: using Run 1 as training	27 out of 29 (93%)	32 out of 35 (91%)	8

# identifying and curating links

## Multiple "John Singer Sargent"

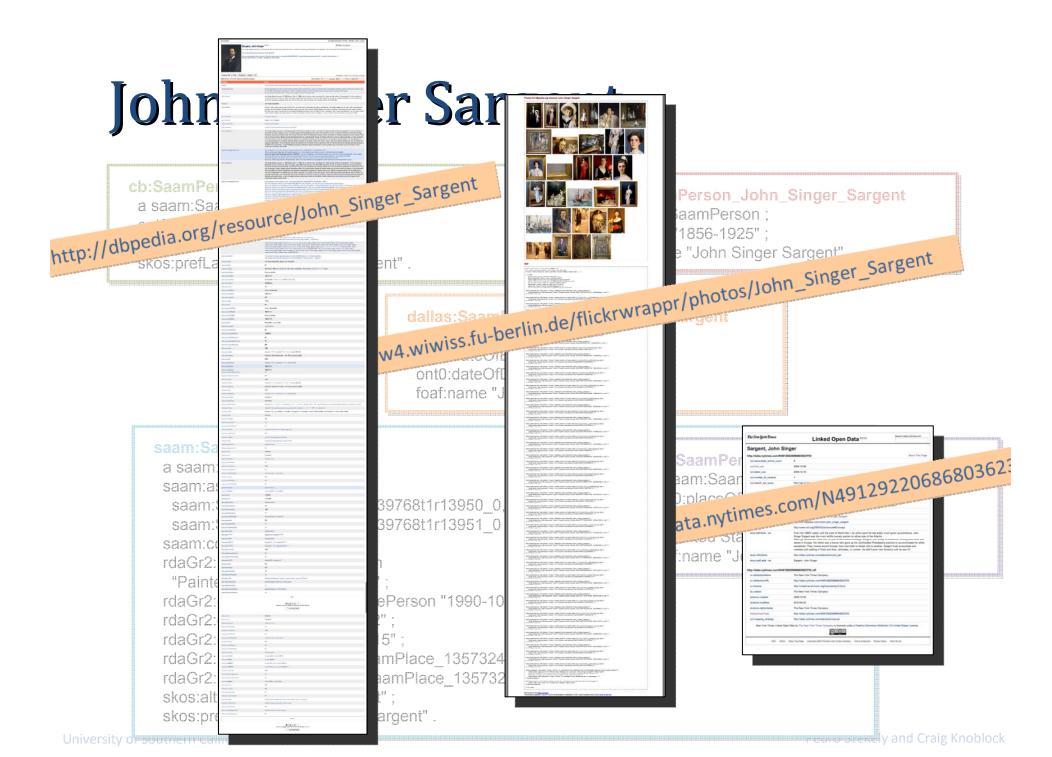
```
cb:Person_John_Singer_Sargent
a aac-ont:Person;
ont0:dateOfBirth "1879", "1885";
ont0:dateOfDeath "1925";
foaf:name "John Singer Sargent".
```

University of

```
ima:Person_John_Singer_Sargent
a aac-ont:Person;
dct:date "1856-1925";
foaf:name "John Singer Sargent".
```

```
dallas:Person_John_Singer_Sargent
a aac-ont:Person;
ont0:dateOfBirth "1856";
ont0:dateOfDeath "1925";
foaf:name "John Singer Sargent".
```

```
saam:Person 4253
                                                             met:Person_John_Singer_Sargent
a aac-ont:Person:
                                                              a aac-ont:Person;
 aac-ont:associatedPlace
                                                              ont0:placeOfResidence
  saam:SaamPlace 1357324439768t1r13950 0,
                                                                "North and Central America".
  saam:SaamPlace 1357324439768t1r13951 0;
                                                               "United States":
 saam:constituentld "4253";
                                                              foaf:name "John Singer Sargent".
 rdaGr2:biographicalInformation
  "Painter. Sargent traveled ...";
 rdaGr2:dateAssociatedWithThePerson "1990-10-1", "1995-5-8";
 rdaGr2:dateOfBirth "1856-1-12":
 rdaGr2:dateOfDeath "1925-4-15":
 rdaGr2:placeOfBirth saam:SaamPlace 1357324439768t1r13952 0 :
 rdaGr2:placeOfDeath saam:SaamPlace 1357324439768t1r13953 0;
foaf:name "John S. Sargent";
skos:altLabel "John S. Sargent";
skos:prefLabel "John Singer Sargent".
                                                                                               and Craig Knoblock
```



## Linking "John Singer Sargent"

```
saam:Person_4253
owl:sameAs cb:Person_John_Singer_Sargent;
owl:sameAs dallas:Person_John_Singer_Sargent;
owl:sameAs ima:Person_John_Singer_Sargent;
owl:sameAs met:Person_John_Singer_Sargent;
owl:sameAs dbpedia:John_Singer_Sargent;
owl:sameAs nytimes:N49129220686803623753;
owl:sameAs w-flick:John_Singer_Sargent;
...
```

### Intuition

## Estimate discrimination power of properties, e.g., of name, birth and death dates

every combination of dates

birth date	death date	# of people
1800	1820	147
1800	1821	284
1800	1822	213

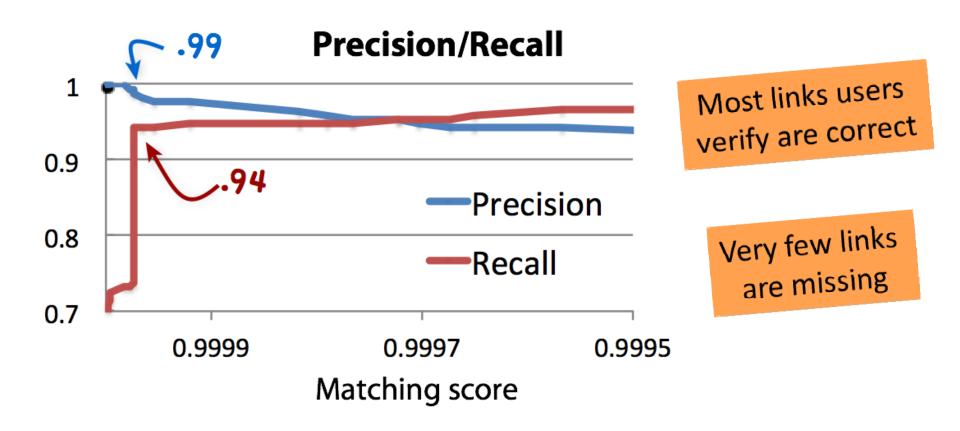
#### similar idea to

Song, D., Heflin, J.: Domain-independent entity coreference for linking ontology instances. ACM Journal of Data and Information Quality (ACM JDIQ) (2012)

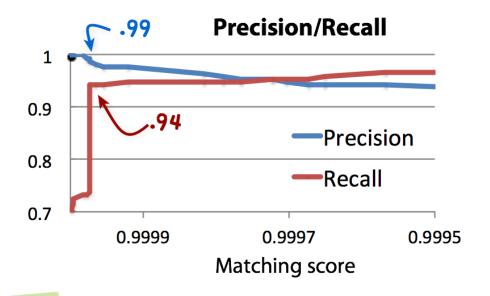
## **Evaluation of Automatic Linking**

SAAM names starting with "A" matched by hand

- **535** people **176** matches



## Results of Automatic Linking



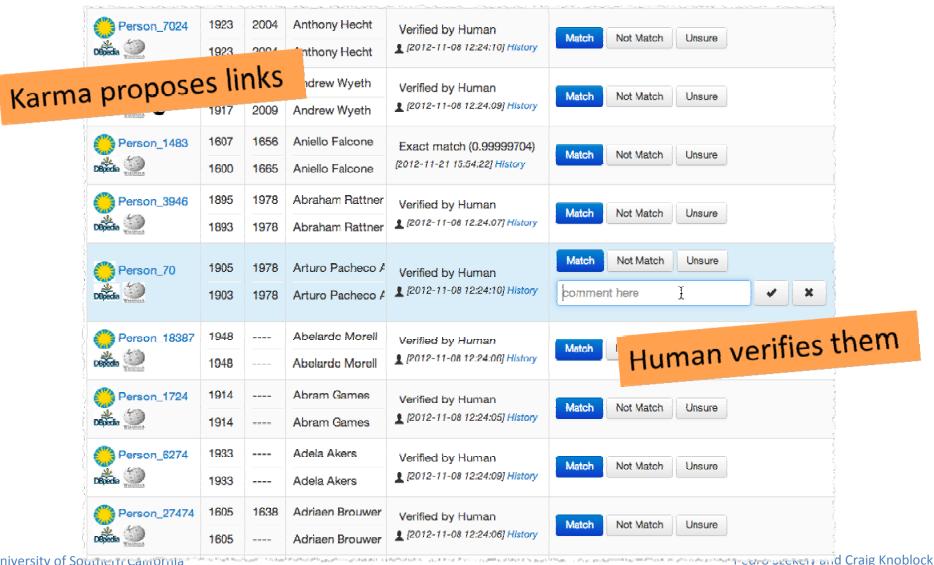
### SAAM verified

DBPedia	2,194
New York Times	70

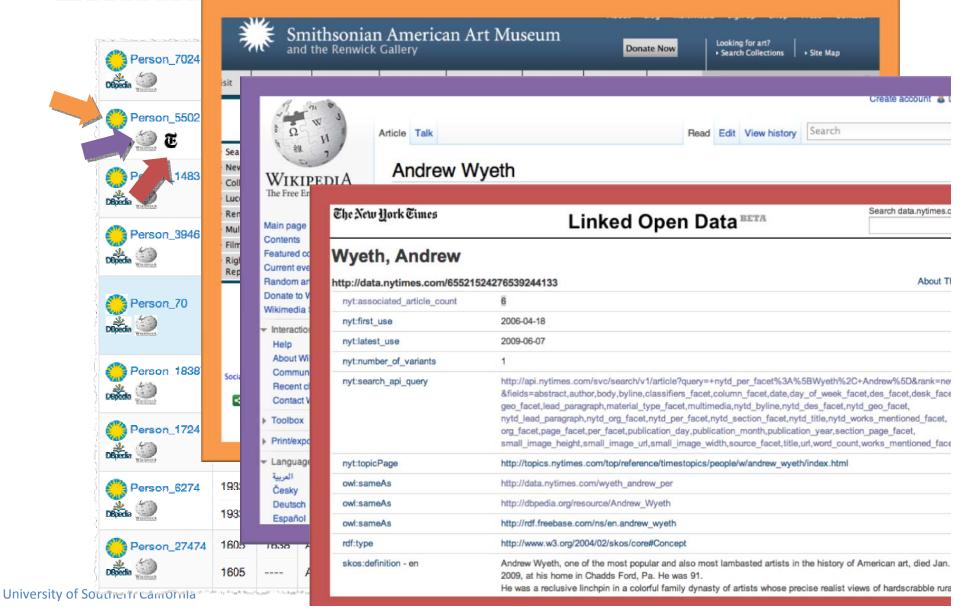
estimate ≈ 30 missing links to DBpedia

Getty ULAN®	2,110
Rijksmuseum	551
Geonames	3,068

## **Curating Links with Karma**



### Linking with Karma



## results of automated linking and interactive curation recorded using PROV

## owl:sameAs statements constructed using SPARQL CONSTRUCT queries over PROV records

## deployment





## public SPARQL endpoint

## Property aac-ont:associatedPlace acc:place/01e03f4e75bd99edafbbf4c4d121646e acc:place/3acced0cccdd4b82ff3dce2e12899609 acc:place/3c2f5fc3870423287905df8210a6c708 acc:place/db4d6d43295c0cbd5be61cd2730f3a3e ElementsGr2:biographicalInformation Painter. Sargent traveled in a circle of socially prominent people and is known for his loosely painted portraits done in a style reminiscent of Edgar Degas and James Abbott McNeill Whistler. Madame X caused a minor scandal at the Salon of 1884 and was rejected by Sargent's client because Sargent depicted her as vain. He devoted his later career

to impressionistic watercolor scenes.

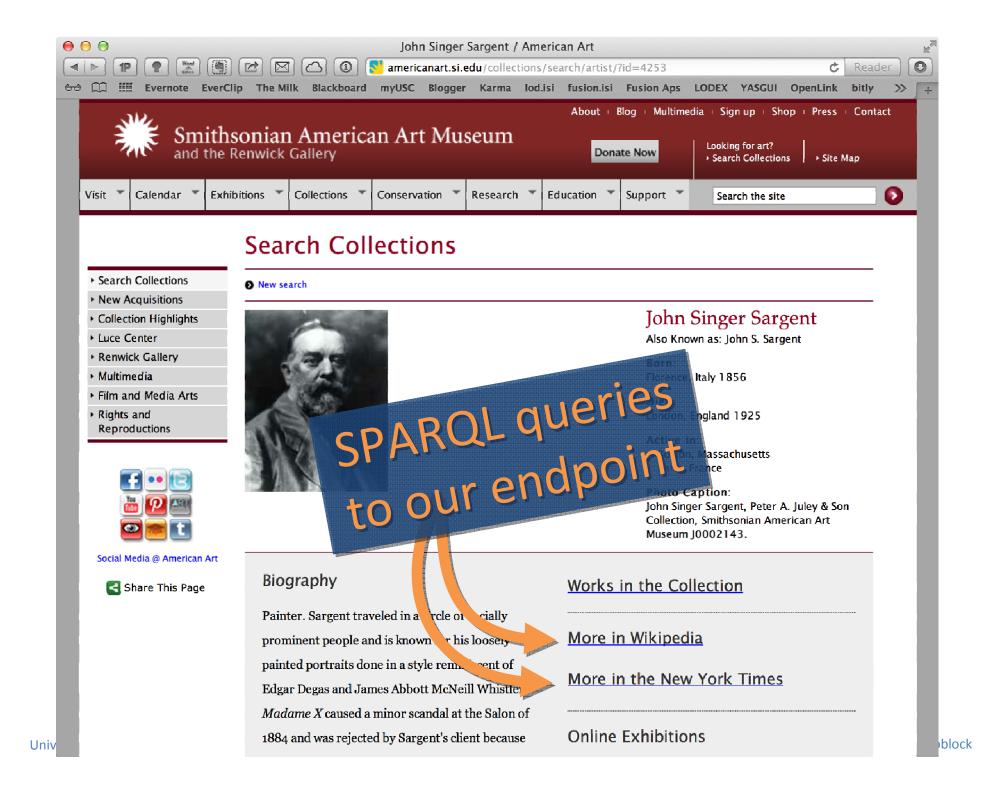
Joan Stahl American Artists in Photographic Portraits from the Peter A. Juley & Son Collection (Washington, D.C. and Mineola, New York: National Museum of American Art and Dover Publications, Inc., 1995)

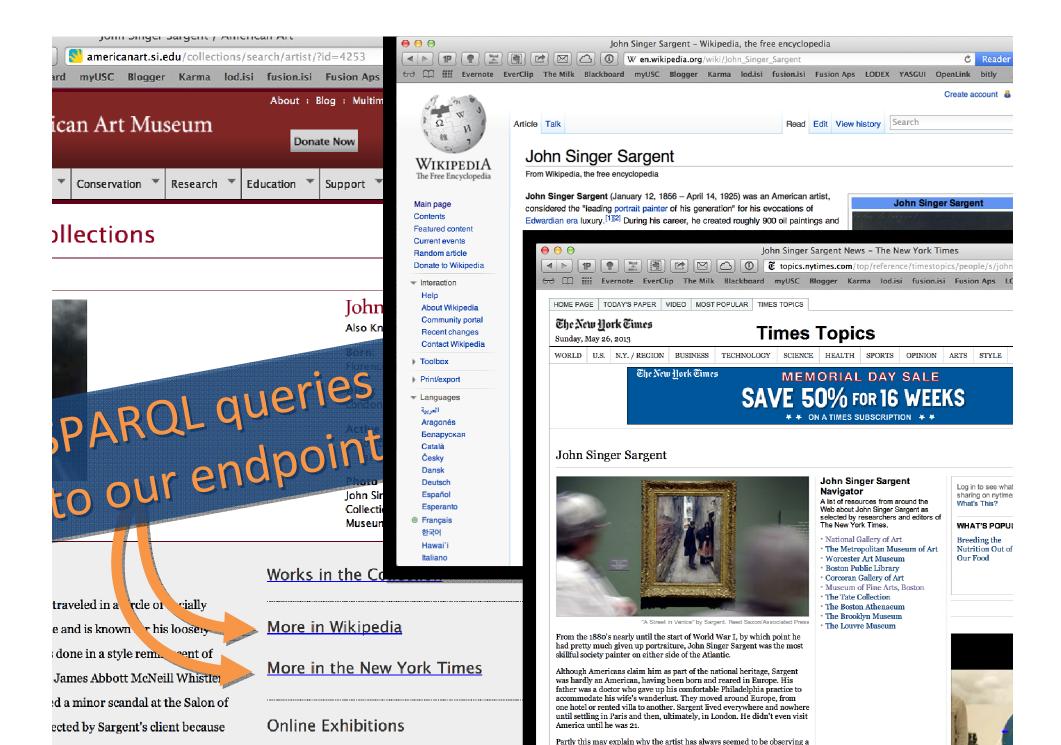
Sargent, born in Italy to American parents and educated in Germany and France, possessed a keen eye for regional
social conventions. As the premier society portrait painter of his day, Sargent instinctively knew what pose, attire,
and mood to suggest in London, New York, or Roston, His English subjects have an air of pomp and splendor. New

	Judith K. Maxwell, Stephen Nissenbaum, Bruce Robertson, Roger B. Stein, and William H. Truettner <i>Picturing Old New England: Image and Memory</i> (Washington, D.C.; New Haven, Conn; and London: National Museum of Americal Art with Yale University Press, 1999)	
saam-ont:constituentId	• 4253	
is dcterms:creator of	aac:object/21033 aac:object/21621 aac:object/21622 aac:object/21623 aac:object/21624 aac:object/28009 aac:object/5151	
aac-ont:creatorOf	<ul> <li>aac:object/21033</li> <li>aac:object/21621</li> <li>aac:object/21622</li> <li>aac:object/21623</li> <li>aac:object/21624</li> <li>aac:object/28009</li> <li>aac:object/5151</li> </ul>	
ElementsGr2:dateAssociatedWithThePe	ersom 1990-01 = 1995-05-08	
ElementsGr2:dateOfBirth	■ 1856-01-12	
ElementsGr2:dateOfDeath	■ 1925-04-15	
foaf:depiction	<a href="http://upload.wikimedia.org/wikipedia/commons/6/6b/John_Singer_Sargentautoportrait_1906.jpg">http://upload.wikimedia.org/wikipedia/commons/6/6b/John_Singer_Sargentautoportrait_1906.jpg</a>	
dc:identifier	4253	
rdfs:label	John Singer Sargent	
foaf:name	<ul> <li>John Singer Sargent</li> </ul>	
ElementsGr2:placeOfBirth	<ul><li>aac:place/3acced0cccdd4b82ff3dce2e12899609</li></ul>	
ElementsGr2:placeOfDeath	<ul> <li>aac:place/3c2f5fc3870423287905df8210a6c708</li> </ul>	
skos:prefLabel	John Singer Sargent	
owl:sameAs	<ul> <li><a href="http://data.nytimes.com/N49129220686803623753">http://dbpedia.org/resource/John_Singer_Sargent&gt;</a></li> <li>:person/4253</li> <li><a href="http://rdf.freebase.com/ns/en.john_singer_sargent">http://rdf.freebase.com/ns/en.john_singer_sargent</a></li> </ul>	
rdfs:seeAlso	<ul><li><a href="http://data.nytimes.com/N49129220686803623753">http://en.wikipedia.org/wiki/John_Singer_Sargent&gt;</a></li></ul>	
saam-ont:seeAlsoNytimes	<http: data.nytimes.com="" n49129220686803623753=""></http:>	
saam-ont:seeAlsoWikipedia	<http: en.wikipedia.org="" john_singer_sargent="" wiki=""></http:>	
rdf:type	<ul> <li>frbr:Person</li> <li>aac-ont:Person</li> <li>edm:Agent</li> <li>owl:Thing</li> <li>skos:Concept</li> <li>foaf:Person</li> </ul>	
aac-ont:variantName	John S. Sargent	

This page shows information obtained from the SPARQL endpoint at http://localhost/openrdf-sesame/repositories/saam-lod.

<u>As Turtle</u> I <u>As RDF/XML</u> I <u>Browse in Disco</u> I <u>Browse in Tabulator</u> I <u>Browse in OpenLink Browser</u>





world he neither occupied nor chose to judge. His pictures are, in a sense,

OTHER STREET OF SOUTH CONTORNIA

### Related Work

- Europeana
  - 17 million items, 1,500 institutions
  - Require exports in "Europeana" format
- Amsterdam Museum, Museum Finland
  - Rich ontology, RDF to RDF mapping rules
- LODAC museums in Japan
  - 114 museums, simple ontology
- Research Space, British Museum
  - CIDOC CRM ontologies, complex mappings

We focused significantly on Linking identification and curation

### **Next Steps**

- Applications leveraging linked data
  - Virtual museum
  - Tools to create multimedia stories about art
  - Tools to find inconsistencies
  - Feed data to wikidata

 American Art Collective: a linked data consortium of museums



## Merci