Prosopography and Computer Ontologies: towards a formal representation of the 'factoid' model by means of CIDOC-CRM



John Bradley, Michele Pasin

Department of Digital Humanities

Kings College, London

michele.pasin@kcl.ac.uk

Summary

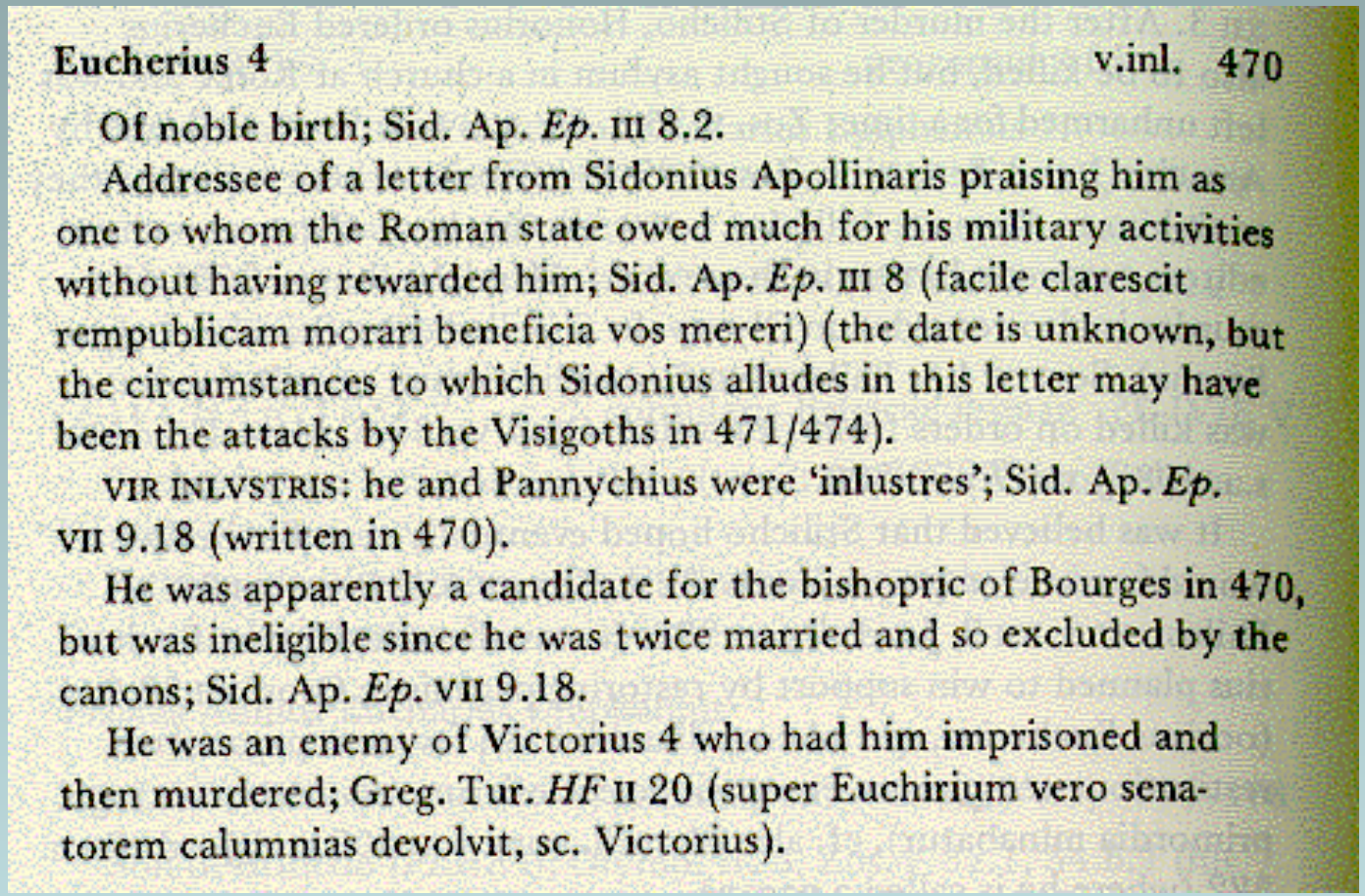
1. **Background**. Structured prosopography; factoid-based pros.; different models for different contexts.

2. **Problem**. Semantic interoperability; expressing the factoid model in a more general fashion.

3. **Approach**. Applying CIDOC-CRM; Strengths and weaknesses; other approaches and future work

Prosopography: from a ‘narrative’ model…

From J.R. Martindale, *The Prosopography of the Later Roman Empire, 3: A.D. 527-641. Cambridge: Cambridge University Press. 1992.*



Eucherius 4 v.inl. 470

Of noble birth; Sid. Ap. *Ep.* III 8.2.

Addressee of a letter from Sidonius Apollinaris praising him as one to whom the Roman state owed much for his military activities without having rewarded him; Sid. Ap. *Ep.* III 8 (facile clarescit rempublicam morari beneficia vos mereri) (the date is unknown, but the circumstances to which Sidonius alludes in this letter may have been the attacks by the Visigoths in 471/474).

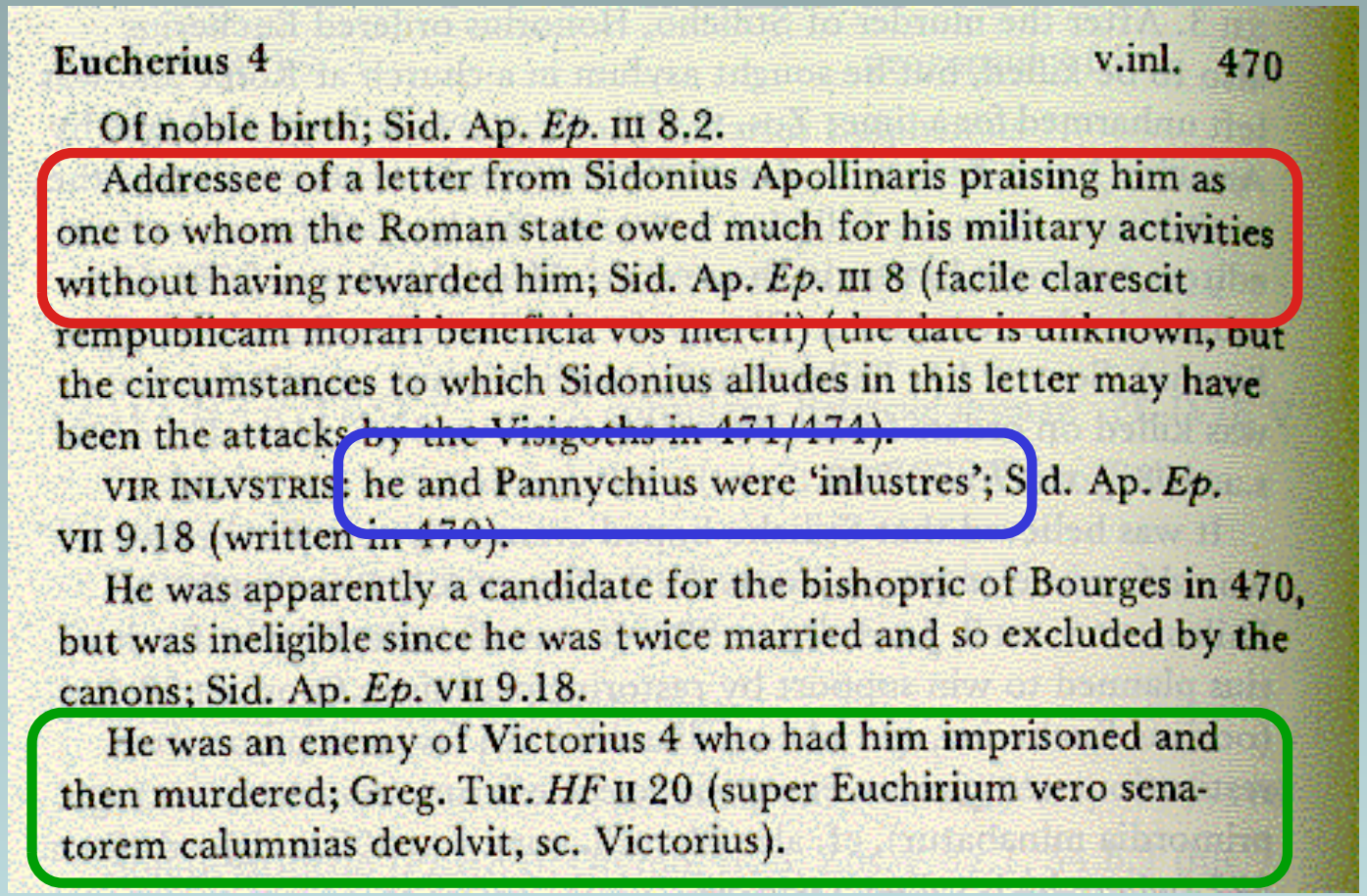
VIR INLVSTRIS: he and Pannychius were ‘inlustres’; Sid. Ap. *Ep.* VII 9.18 (written in 470).

He was apparently a candidate for the bishopric of Bourges in 470, but was ineligible since he was twice married and so excluded by the canons; Sid. Ap. *Ep.* VII 9.18.

He was an enemy of Victorius 4 who had him imprisoned and then murdered; Greg. Tur. *HF* II 20 (super Euchirium vero senatorem calumnias devolvit, sc. Victorius).

Prosopography: … to a ‘structured’ model

- Personal Information: “A was described this way”
- Status: “A held this *status*”
- Event: “A took part in this event of this *type*”



Eucherius 4 v.inl. 470

Of noble birth; Sid. Ap. *Ep.* III 8.2.

Addressee of a letter from Sidonius Apollinaris praising him as one to whom the Roman state owed much for his military activities without having rewarded him; Sid. Ap. *Ep.* III 8 (facile clarescit rempublicam morari beneficia vos inereri) (the date is unknown, but the circumstances to which Sidonius alludes in this letter may have been the attacks by the Visigoths in 471/474).

VIR INLVSTRIS: he and Pannychius were ‘inlustres’; Sid. Ap. *Ep.* VII 9.18 (written in 470).

He was apparently a candidate for the bishopric of Bourges in 470, but was ineligible since he was twice married and so excluded by the canons; Sid. Ap. *Ep.* VII 9.18.

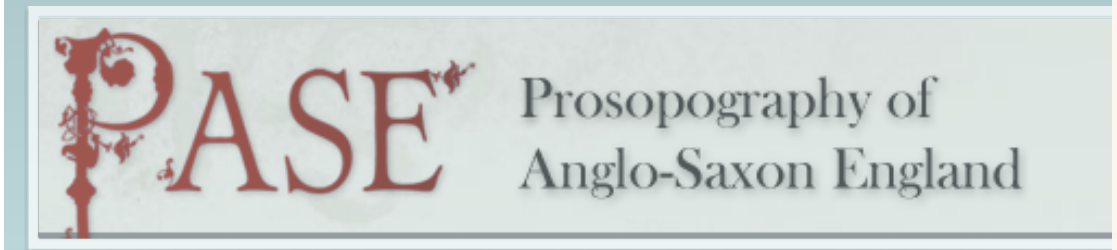
He was an enemy of Victorius 4 who had him imprisoned and then murdered; Greg. Tur. *HF* II 20 (super Euchirium vero senatorem calumnias devolvit, sc. Victorius).

Structured Prosopography: information types











- **Authorship**: "A authored text B"
- **Education**: "A was educated by B"
- **Event**: "A took part in this event of this *type*"
- **Personal Relationship**: "A was related to B (e.g. sister)"
- **Occupation**: "A practiced this *occupation*"
- **Office**: "A held this *office*"
- **Personal Information**: "A was described this way (e.g. saintly)"
- **Possession**: "A owned this thing"
- **Recorded Name**: "A's name was written this way"
- **Status**: "A held this *status*"
- **Transaction**: "A took part in this transaction"

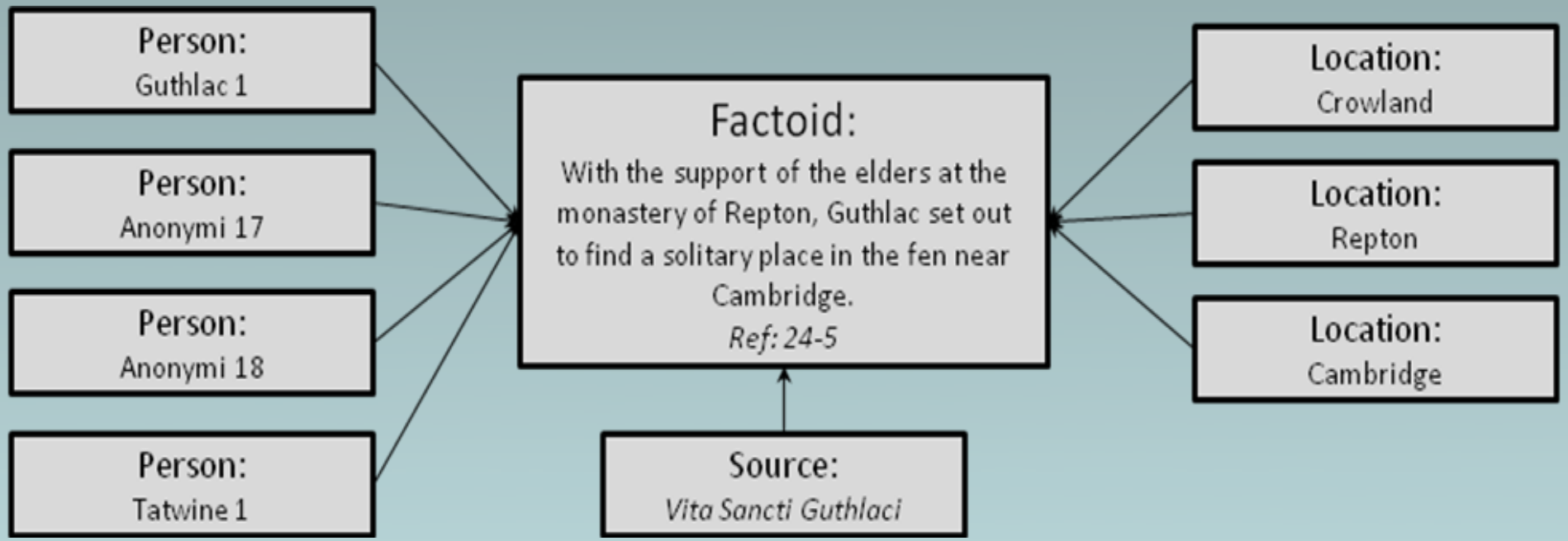
But…

“So, if the king could not fully bring his sense of and desire for order to Scottish society, how are modern historians supposed to do so? To what extent can a highly formal structure such as one finds in a database be useful or helpful to represent the complex and ambiguous aspects of this society?”

Fact vs Factoid

- A Factoid is
  - An assertion made by the project team that **a source "S" at reference "R" states something ("F") about a person or persons ("P")**
- A "Fact" (Boute 2002)
  - "something that happens to a person at a certain moment"
    - Boute, Bruno (2002). Towards More Uniform Database Structures for Prosopographical Research: Work in Progress in University History -- the Example of the Lovanienses Database. In Keats-Rohan, K.S.B. (ed.) *Resourcing Sources*. Oxford: Unit for Prosopographical Research.

Factoid model: an abstract diagram



- notice the ‘gap’ between sources & assertions

Factoid model: how they look for users

Permalink: poms.cch.kcl.ac.uk/db/record/person/762/

Viewing Person Record: Adam of Eccles, master (fl.1178×98)

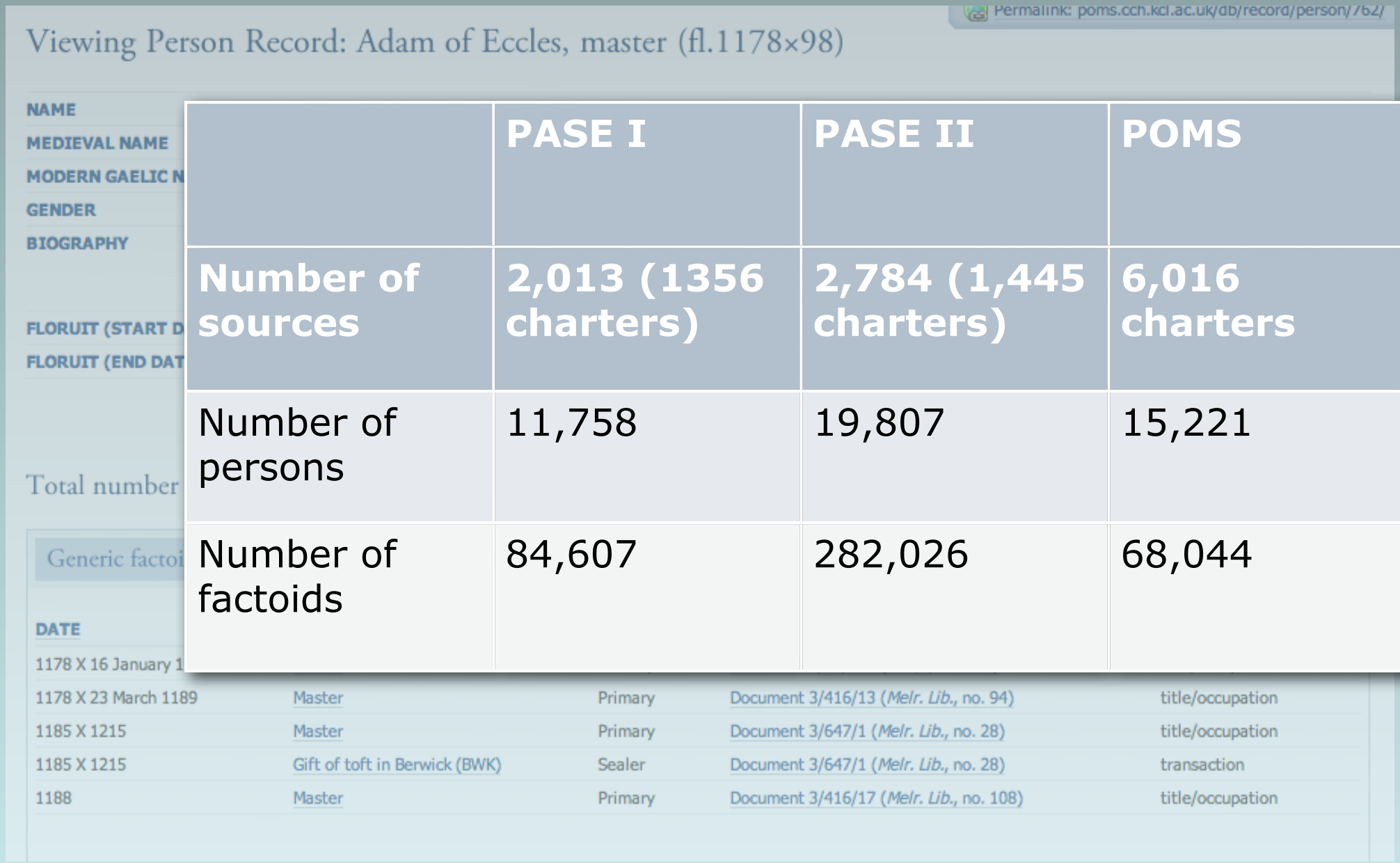
NAME	Adam of Eccles, master (fl.1178×98)
MEDIEVAL NAME	Ádam
MODERN GAELIC NAME	Ádhamh
GENDER	M
BIOGRAPHY	Watt, _Graduates_, 170: Adam was probably from Eccles (BWK) or held the church there, and which may have been held after him by Henry of Eccles, before passing in 1209 to Radulf, priest of Dunbar. He is described as 'Adam the priest' on his seal, but is found as Master 1178×88. He appears witnessing charters from 1186×89, and ×1198. He might be the same as Adam de Nichole. Cf. Adam of Hirsel.
FLORUIT (START DATE)	1185
FLORUIT (END DATE)	1188

Total number of associated factoids: 9

Generic factoids (5): Witnesses factoids (4):

DATE	SHORT SUMMARY	ROLE	SOURCE	TYPE
1178 X 16 January 1188	Master	Primary	Document 2/10/108 (*SEA*, i, no. 220)	title/occupation
1178 X 23 March 1189	Master	Primary	Document 3/416/13 (*Melr. Lib.*, no. 94)	title/occupation
1185 X 1215	Master	Primary	Document 3/647/1 (*Melr. Lib.*, no. 28)	title/occupation
1185 X 1215	Gift of toft in Berwick (BWK)	Sealer	Document 3/647/1 (*Melr. Lib.*, no. 28)	transaction
1188	Master	Primary	Document 3/416/17 (*Melr. Lib.*, no. 108)	title/occupation

Factoid model: some real numbers



	PASE I	PASE II	POMS
Number of sources	**2,013 (1356 charters)**	**2,784 (1,445 charters)**	**6,016 charters**
Number of persons	11,758	19,807	15,221
Number of factoids	84,607	282,026	68,044

What is the factoid model for?

> “By "modeling" I mean the heuristic process of constructing and manipulating models, a "model" I take to be either a representation of something for purposes of study, or a design for realizing something new.”

McCarty, W. (2004). *Modeling: A Study in Words and Meanings*. In S. Schreibman, R. Siemens, & J. Unsworth (Eds.), A Companion to Digital Humanities. Oxford: Blackwell.

A Quadripartite view of models in DH

Data Acquisition



Workflow, experts habits and conceptualization

Data Storage



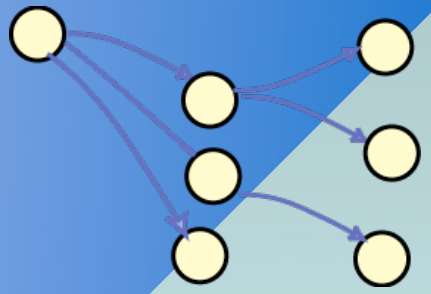
Efficiency, price, platform

Data Presentation



Tables, colors, information layout & organization, discipline-specific discourse

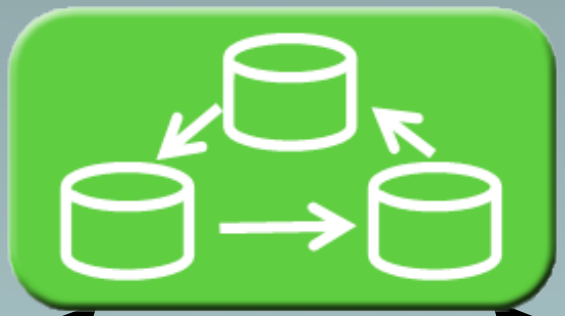
Data Integration



Semantic interoperability, adherence to recognised standards

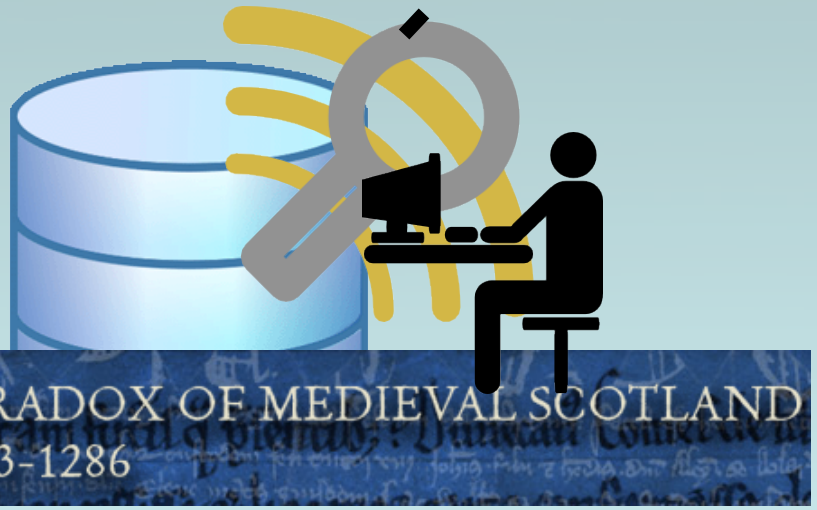
The importance of semantic interoperability

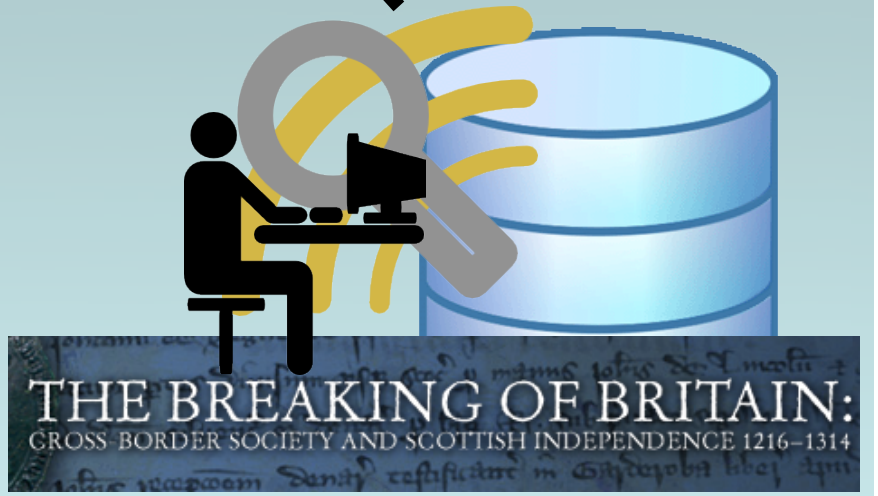




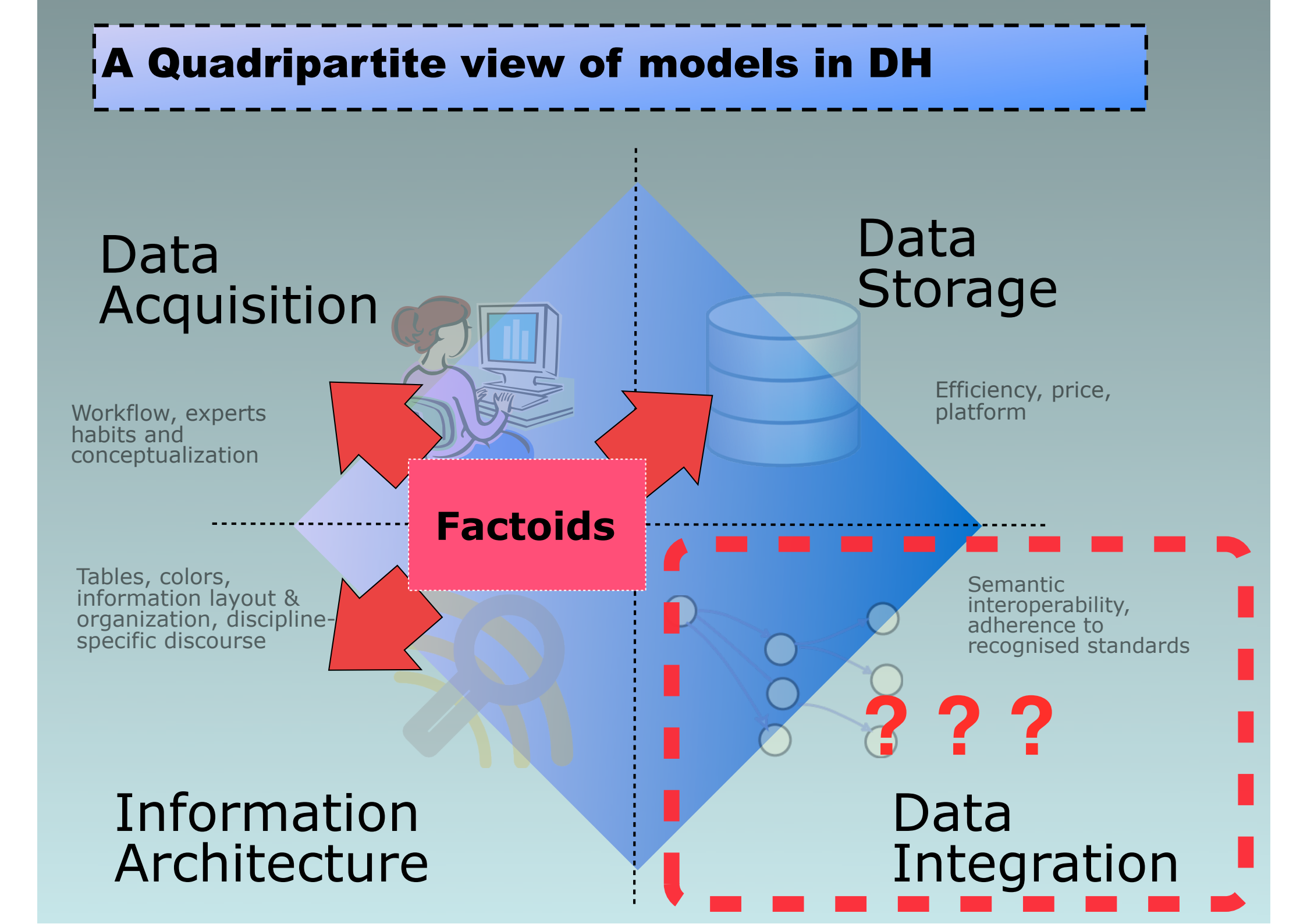
Person:
Adam of Eccles

Person:
Adam of Eccles

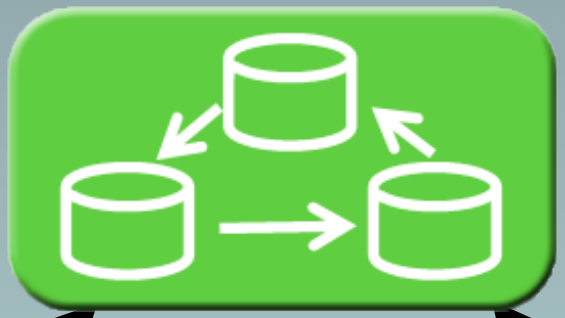




A Quadripartite view of models in DH



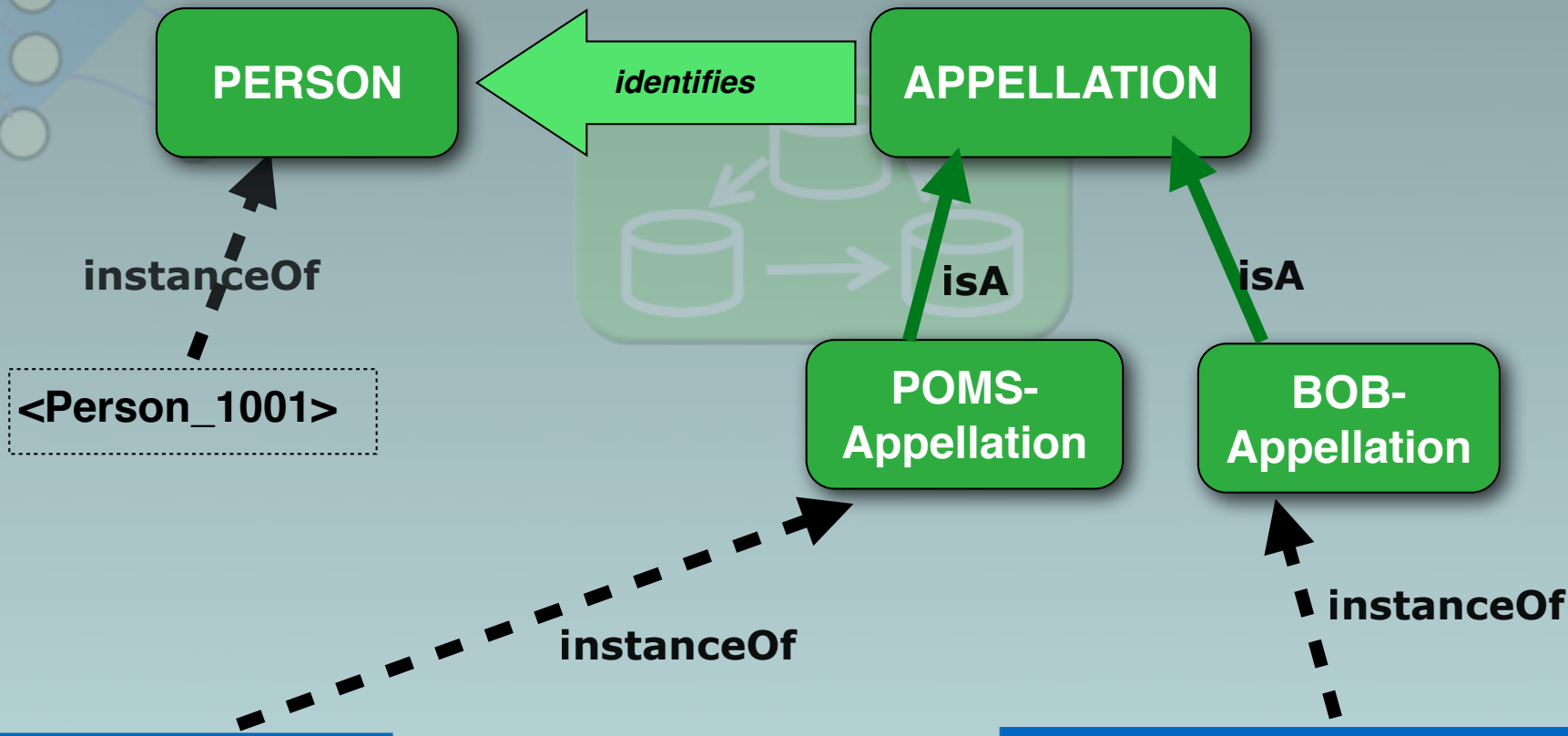
The importance of semantic interoperability



PERSONS *(POMS project)*	
Name	**Surname**
Adam	of Eccles

PERSONS *(BOB project)*	
Name	**Place Name**
Adam of Eccles	Eccles

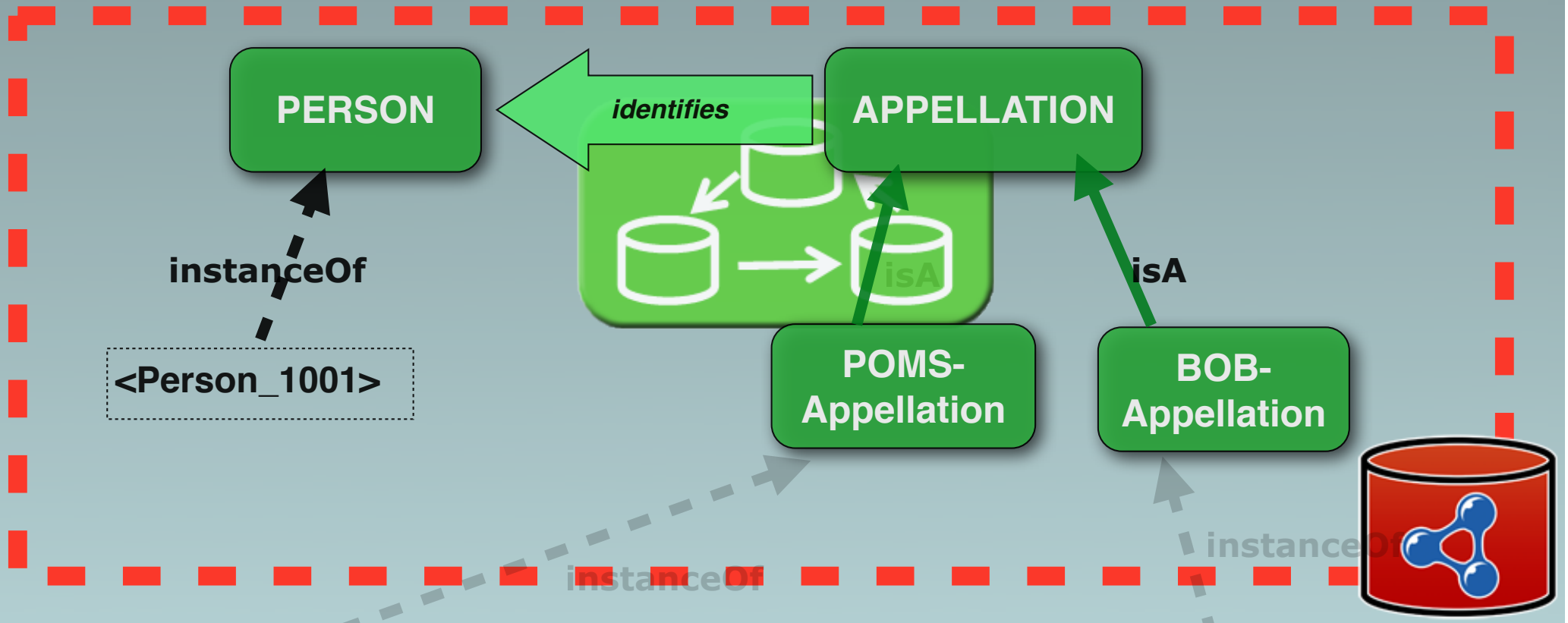
The importance of semantic interoperability



PERSONS *(POMS project)*	
Name	**Surname**
Adam	of Eccles

PERSONS *(BOB project)*	
Name	**Place Name**
Adam of Eccles	Eccles

Ontology-based semantic integration



- formal ontology provides a *theory* of how to make categorical distinctions in *systematic* and *coherent* manner

- making representational choices at the ***highest*** **level of *abstraction*, while still being as clear as possible about the meaning of terms**

Example: the CIDOC-CRM ontology

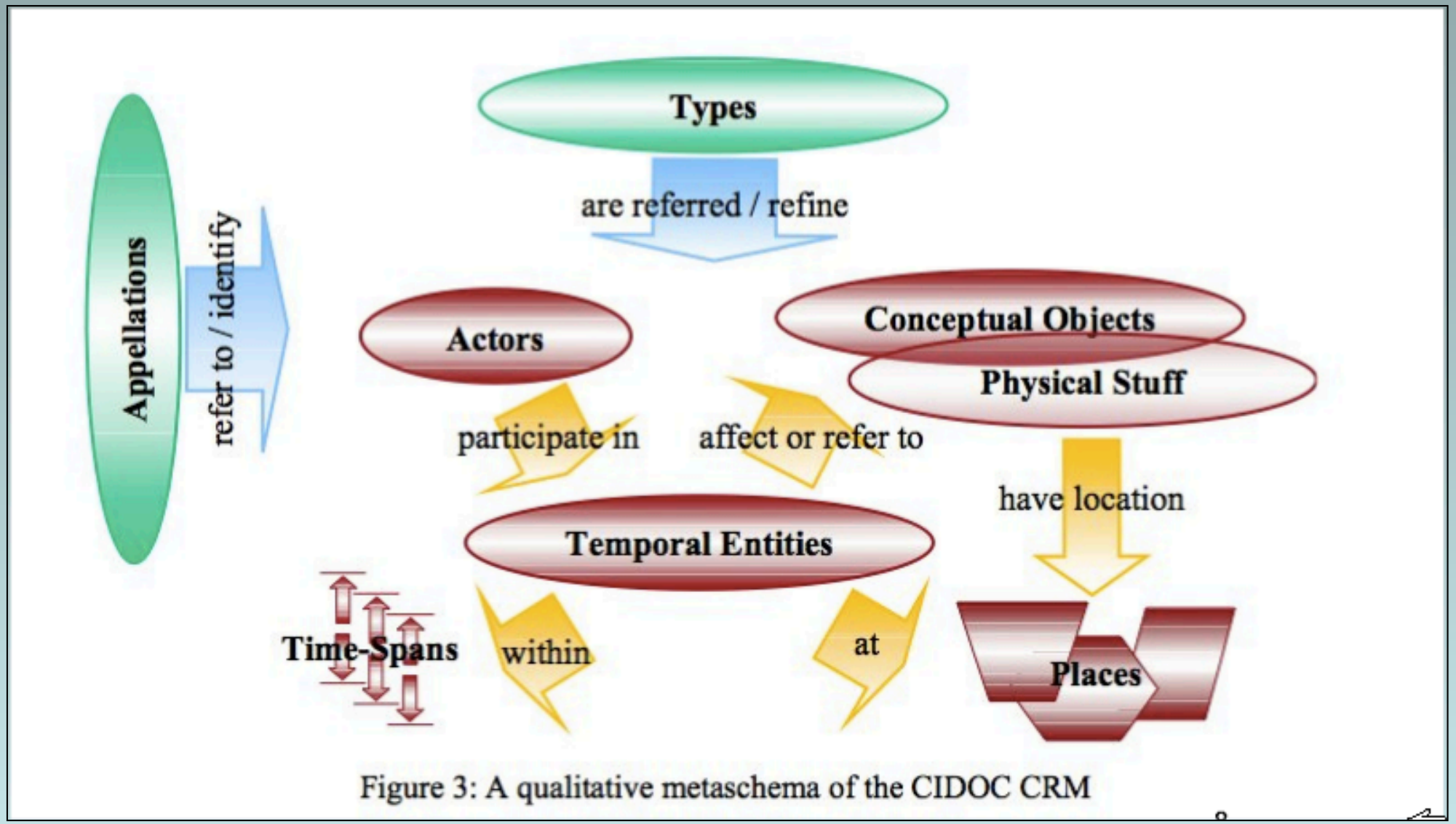
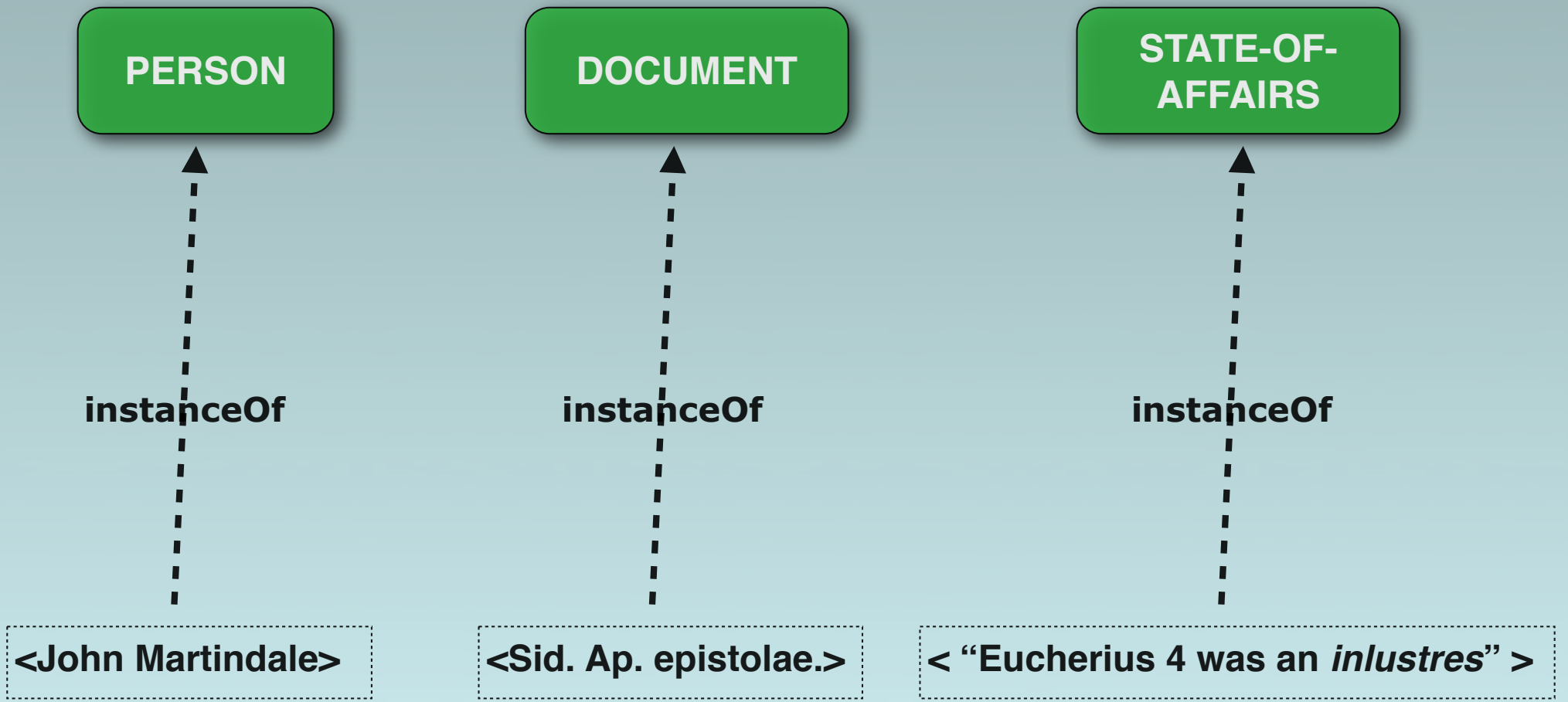
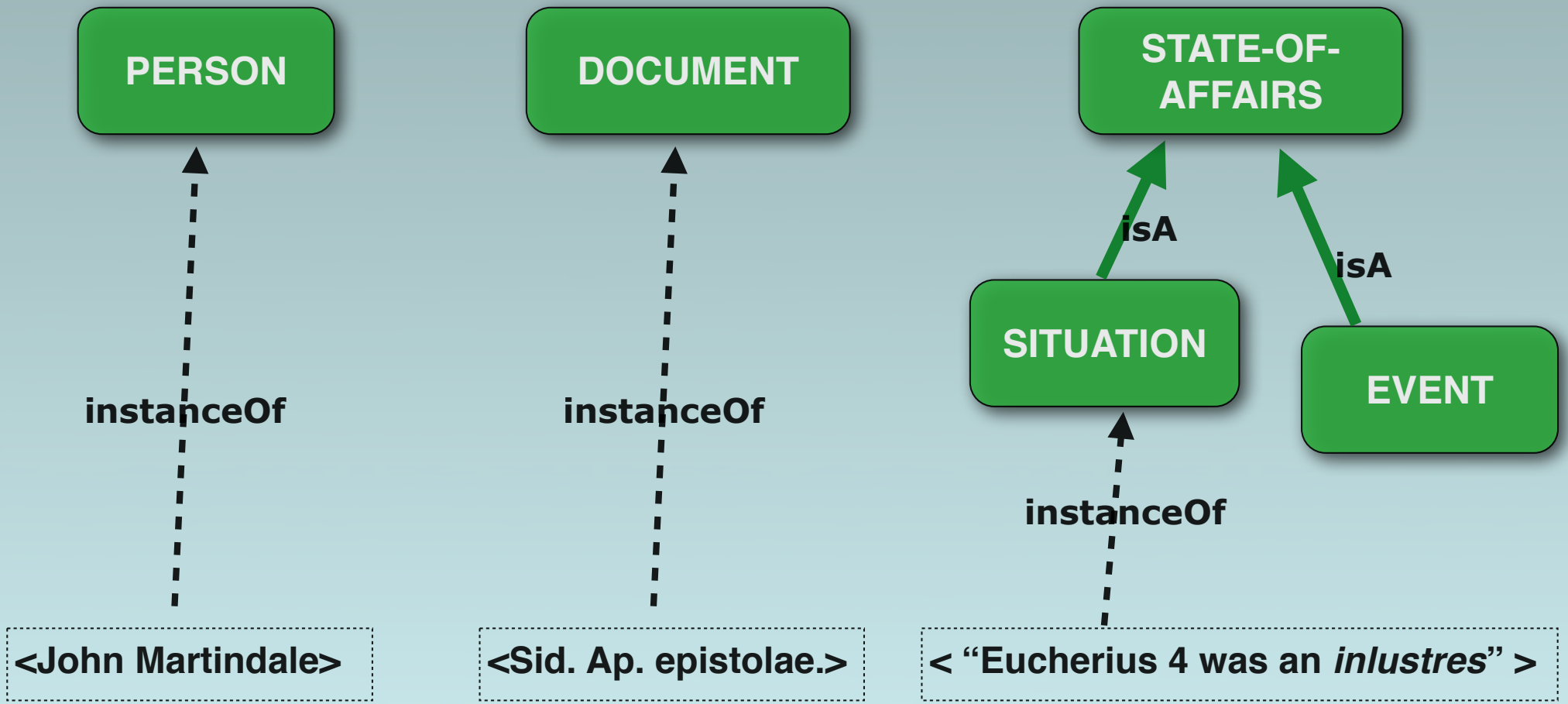


Figure 3: A qualitative metaschema of the CIDOC CRM

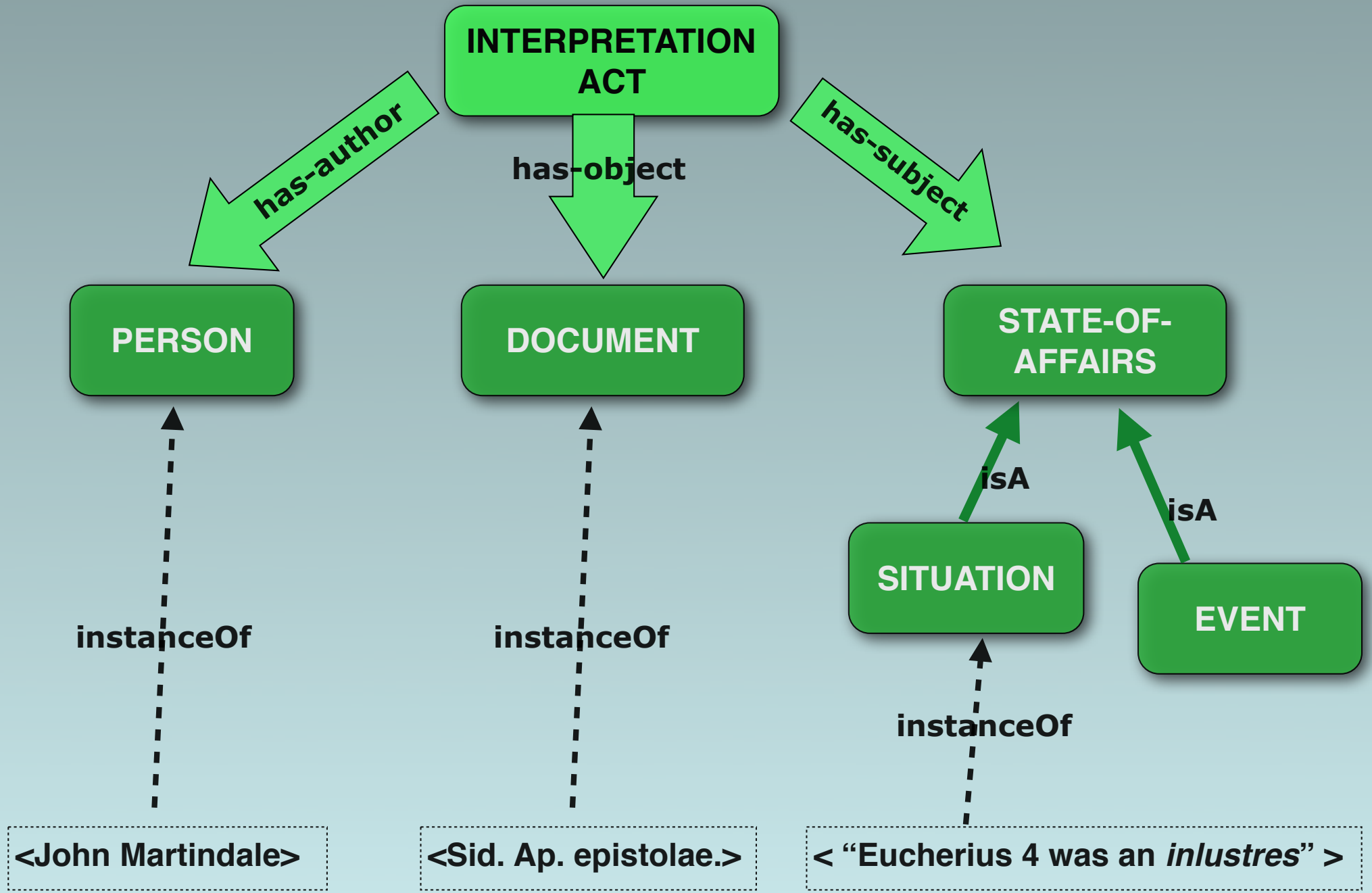
Moving towards a more interoperable model: what are ‘factoids’ in ontological terms?



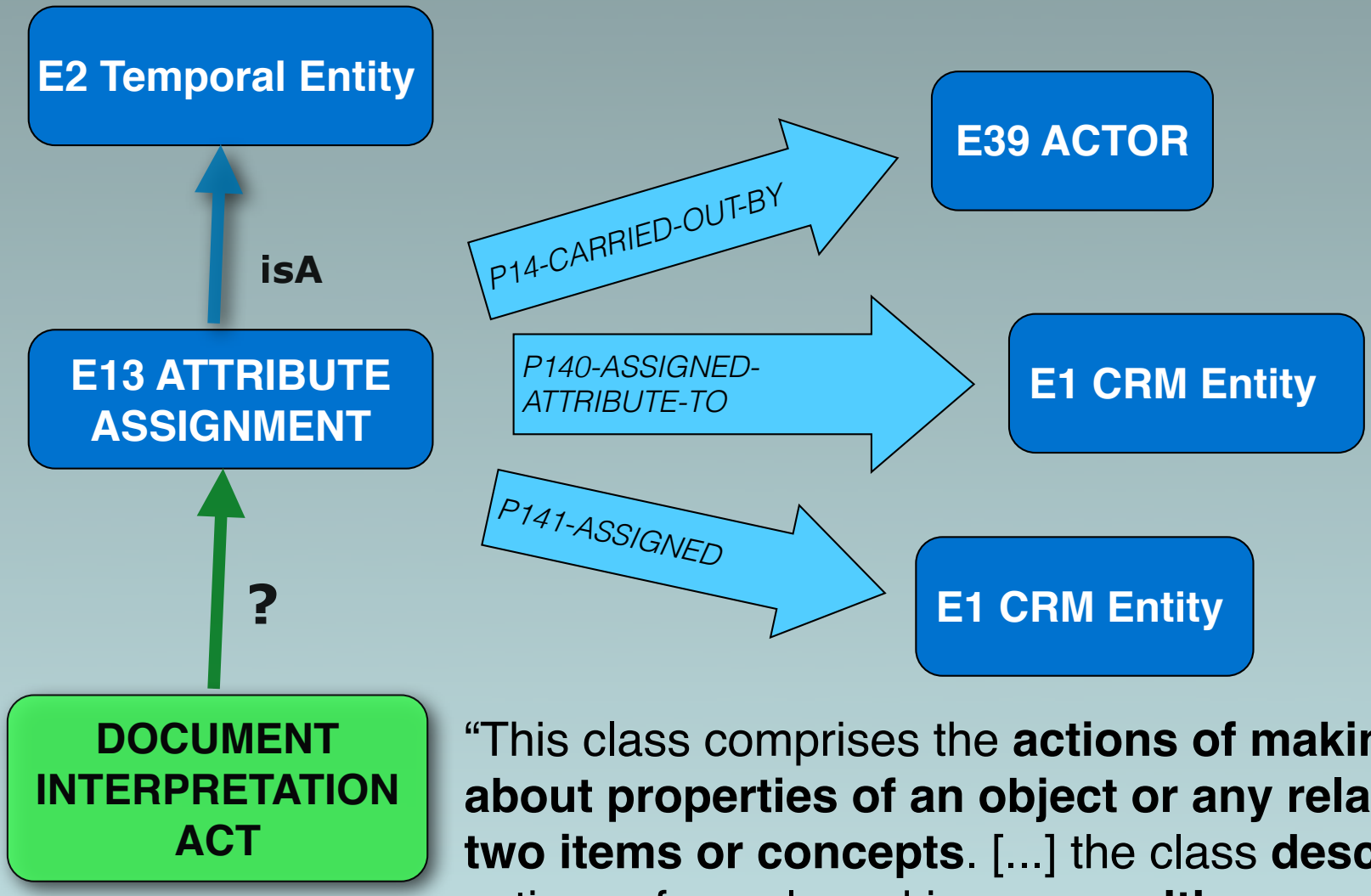
Moving towards a more interoperable model: what are ‘factoids’ in ontological terms?



Factoids as ‘Interpretation’ events

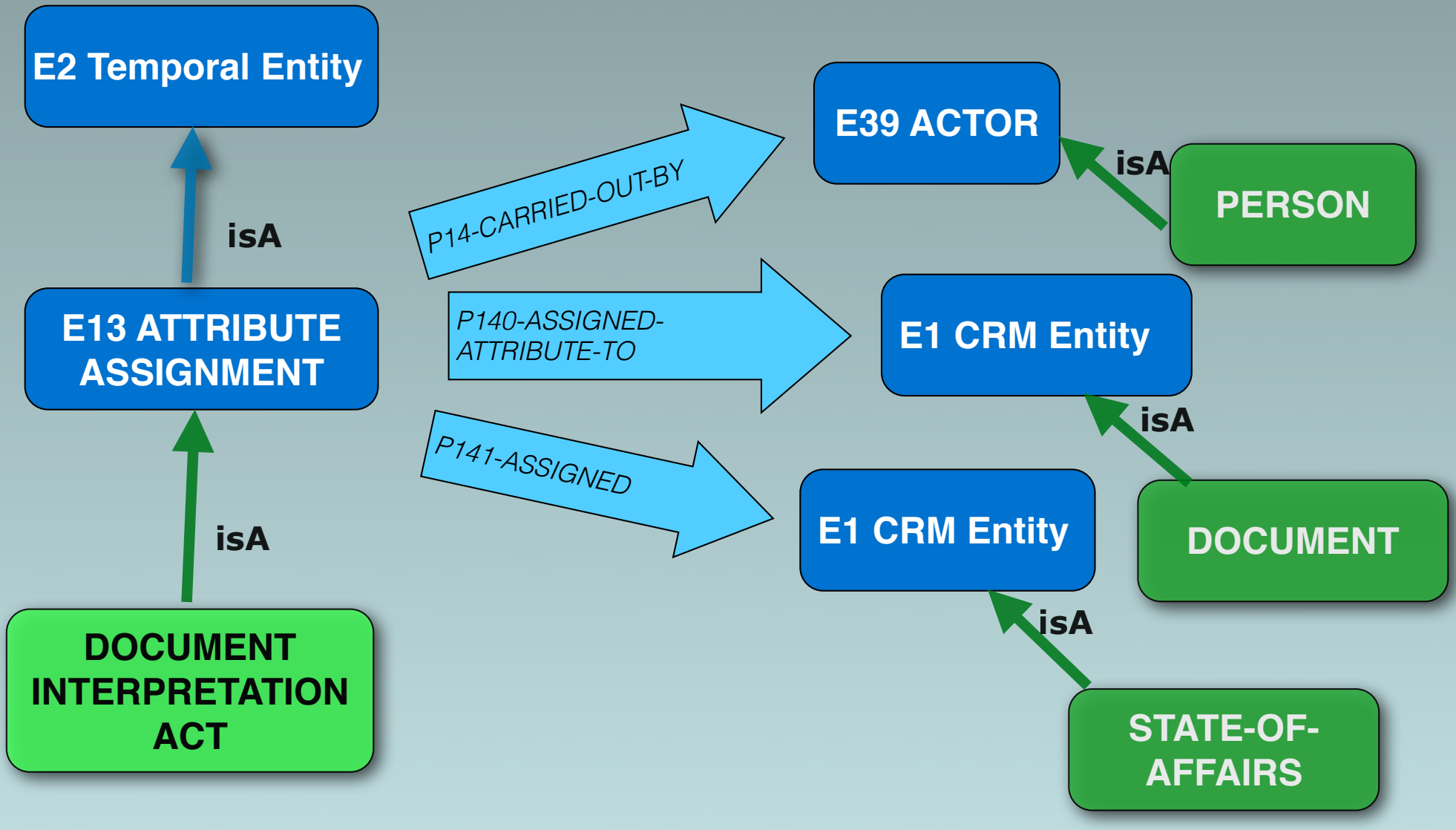


Wrapping things up using CIDOC-CRM



"This class comprises the **actions of making assertions about properties of an object or any relation between two items or concepts**. [...] the class **describes** the actions of people making **propositions and statements during certain museum procedures**, e.g. the person and date when a condition statement was made, an identifier was assigned, the museum object was measured, etc."

Wrapping things up using CIDOC-CRM

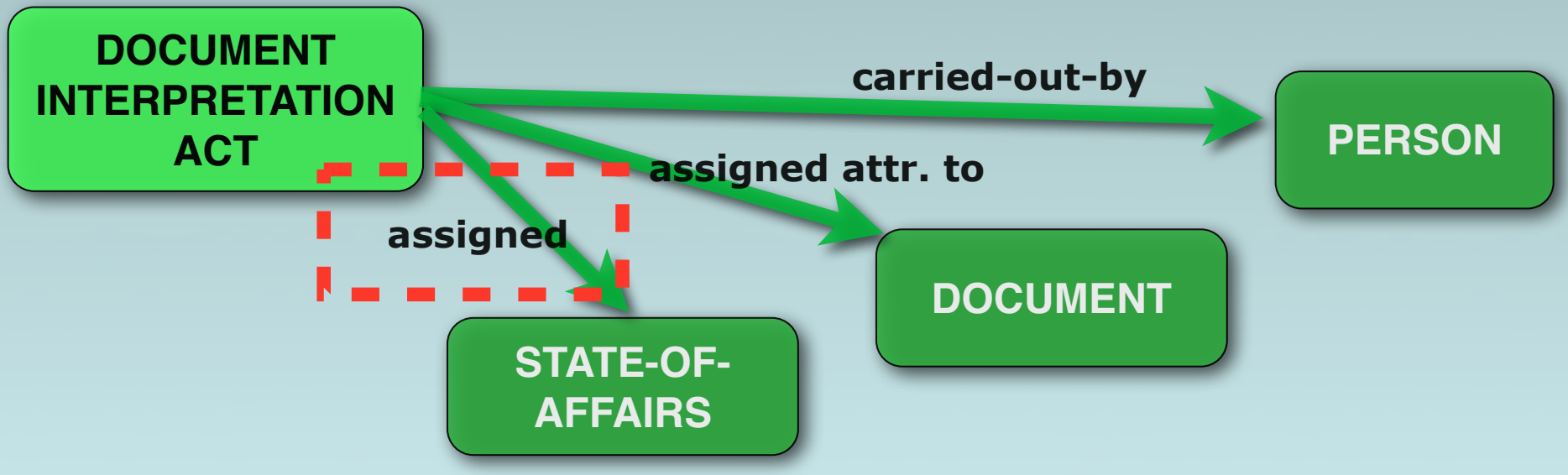


Open issues #1

- the ‘assigned’ property is too generic for rendering our source-based approach to prosopography

-eg. ‘is-evidence-for’; ‘is-about-event’; ‘claims-that’, ‘asserts’

- the key issue derives from the fact that factoids must **reflect** the source as close as possible

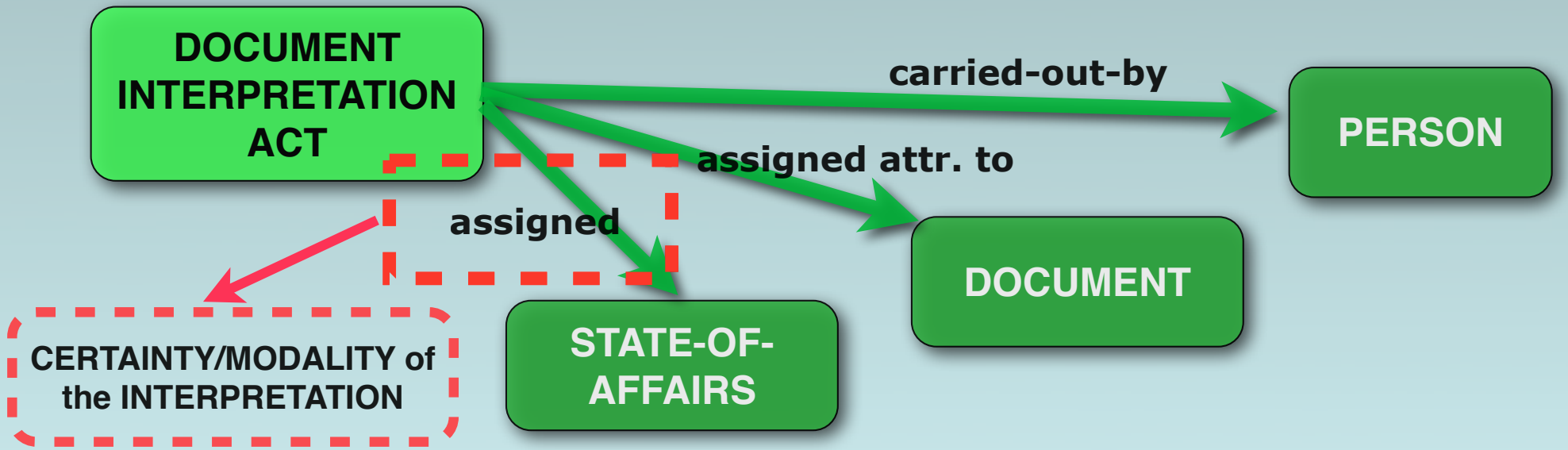


Open issues #2

- if we analyze more seriously the notion of ‘interpretation’ as the institution of a connection between two entities - the connection itself should become a *parameter*!

-eg. We can lay out alternatives in a continuum: ‘is-evidence-for’; ‘is-about-event’; ‘claims-that’ as carrying different pre-defined semantics

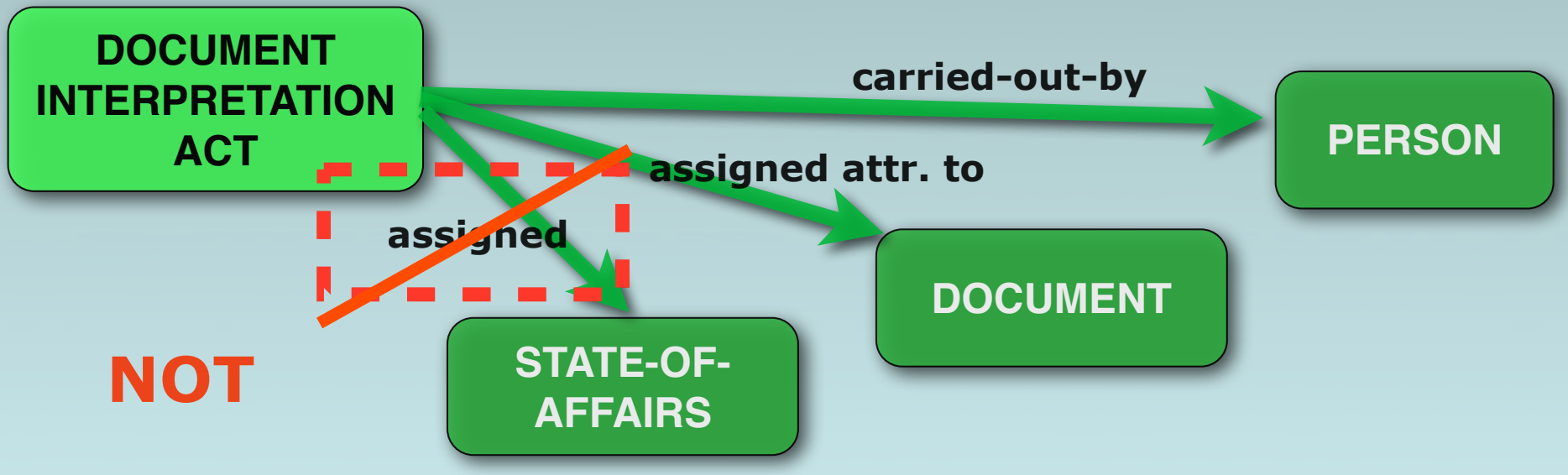
-eg. Opens up interesting paths towards **modal logic:** *X might-be about Y*



Open issues #3

- it seems plausible to create negative ‘interpretation’ - how would this play out in the context of a prosop. database?

-eg. “According to source X, we know that the King of England wasn’t involved in event E”



Summary

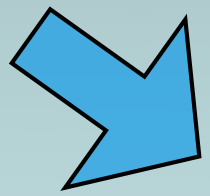
- The factoid-based approach in structured prosopography is a successful one

- It can be ‘ontologized’ using CIDOC-CRM

- Some features need an ad hoc treatment...

- **Next steps**: formalize adequate extensions for CIDOC-CRM & make POMS/BOB datastore accessible via APIs

... thanks!



http://www.poms.ac.uk/

http://www.breakingofbritain.ac.uk/

email me at: michele.pasin@kcl.ac.uk