Core Curriculum in Clinical and Translational Research Seminar Series
Biomedical Informatics Series Part 3 of 4

Pain and Mental Health: A Case Study in Information-Driven Research

February 22, 2012 – Roswell Park Cancer Institute, Buffalo NY

Werner CEUSTERS, MD

Center of Excellence in Bioinformatics and Life Sciences, Ontology Research Group
and
Department of Psychiatry

University at Buffalo, NY, USA

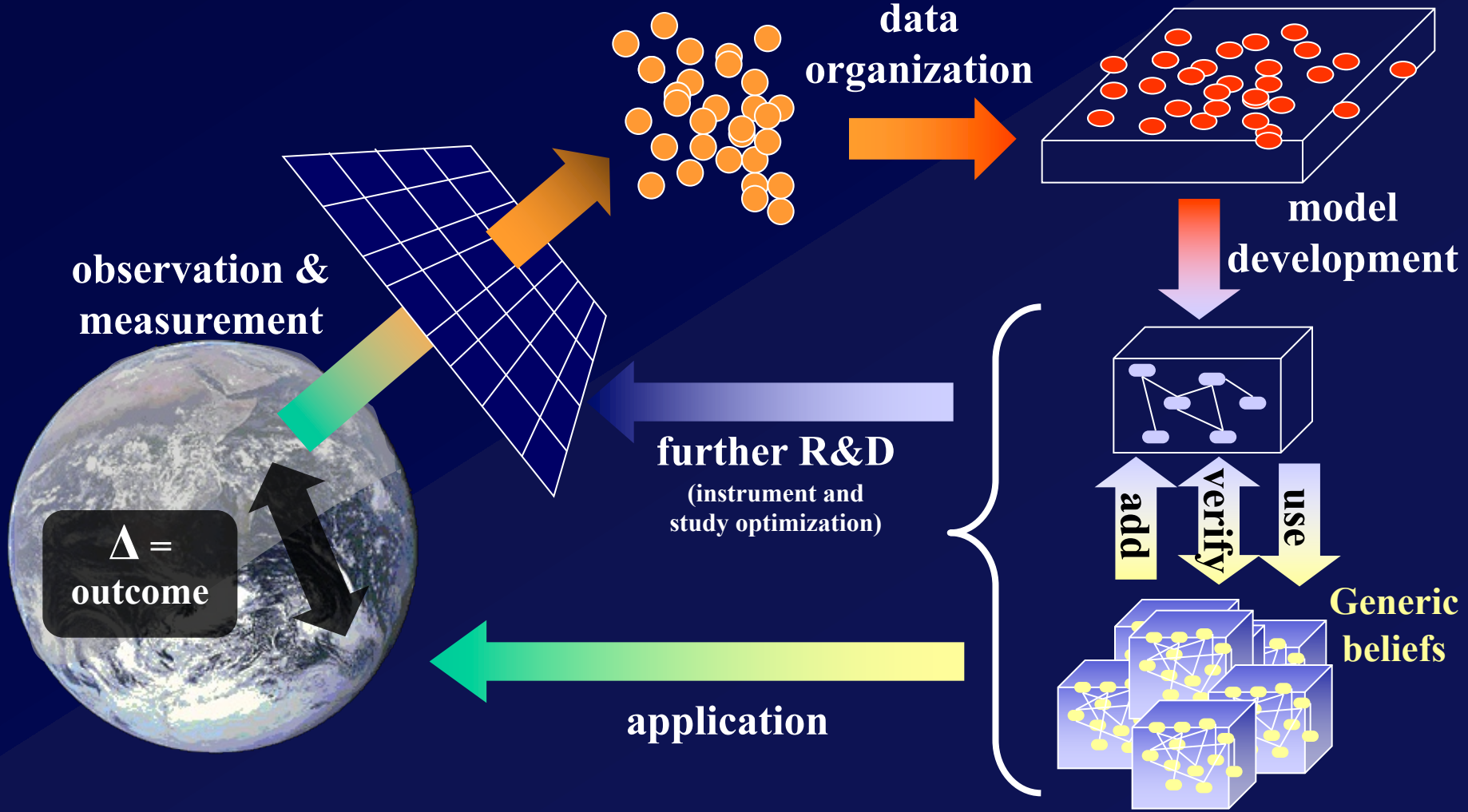
http://www.org.buffalo.edu/RTU

Overview

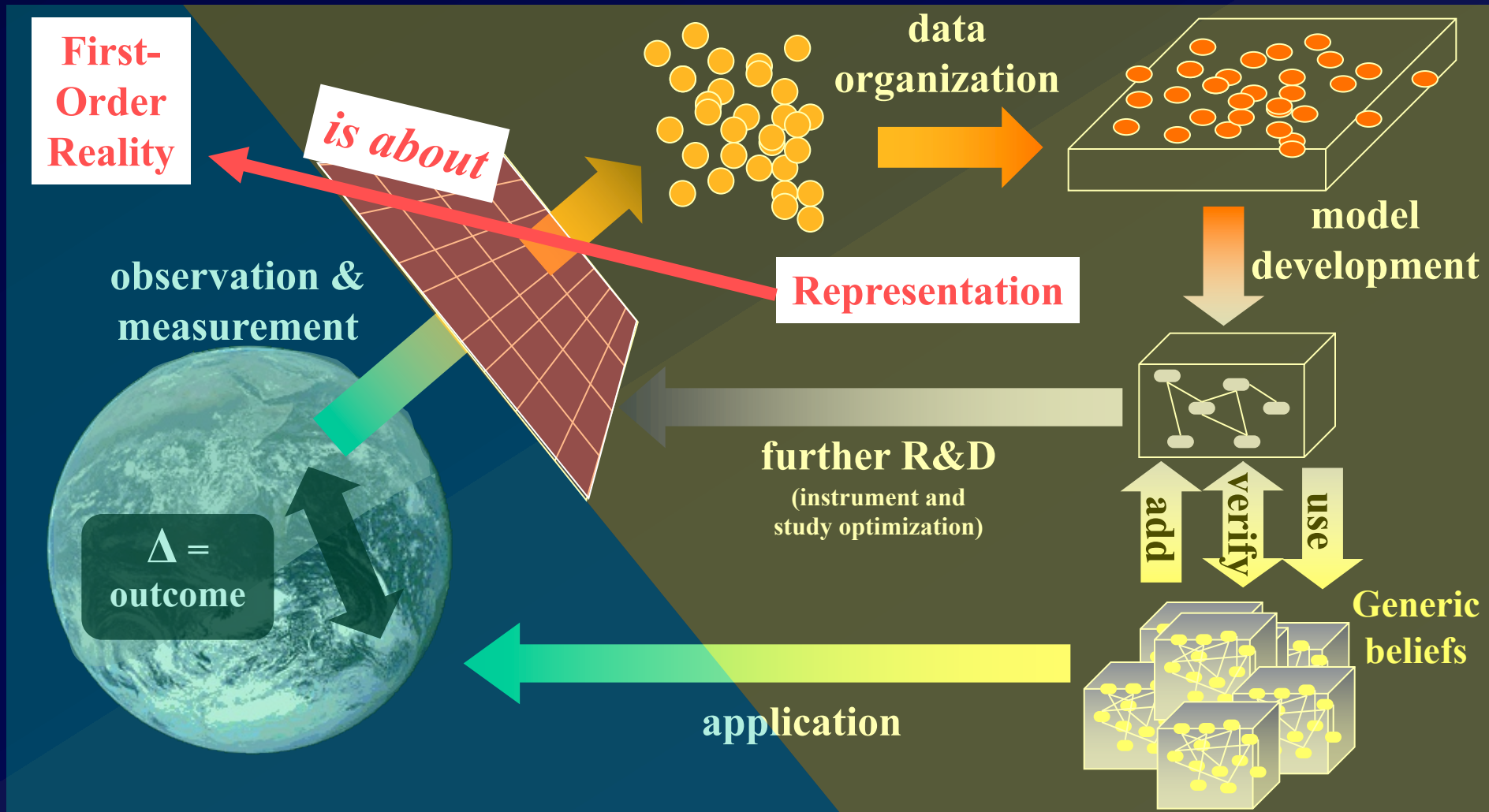
- Principles of data collection
- Introduction to OPMQoL: an ontology for pain-related disablement, mental health and quality of life,
- Making individual data collections more useful for international research,
- Introduction to application ontology building in for OPMQoL.

Principles of Data Collection

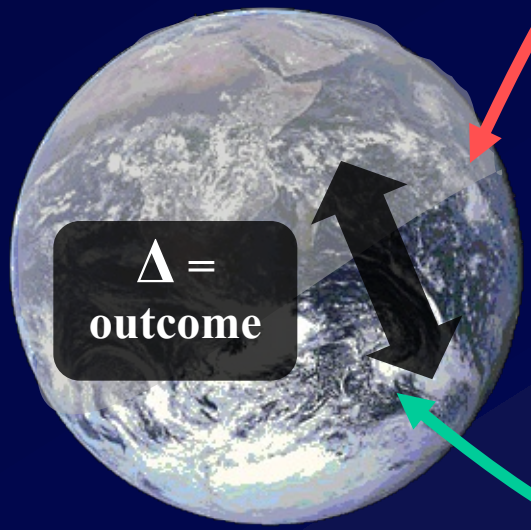
Data generation and use



A crucial distinction: data and what they are *about*



Data must be unambiguous and faithful to reality …

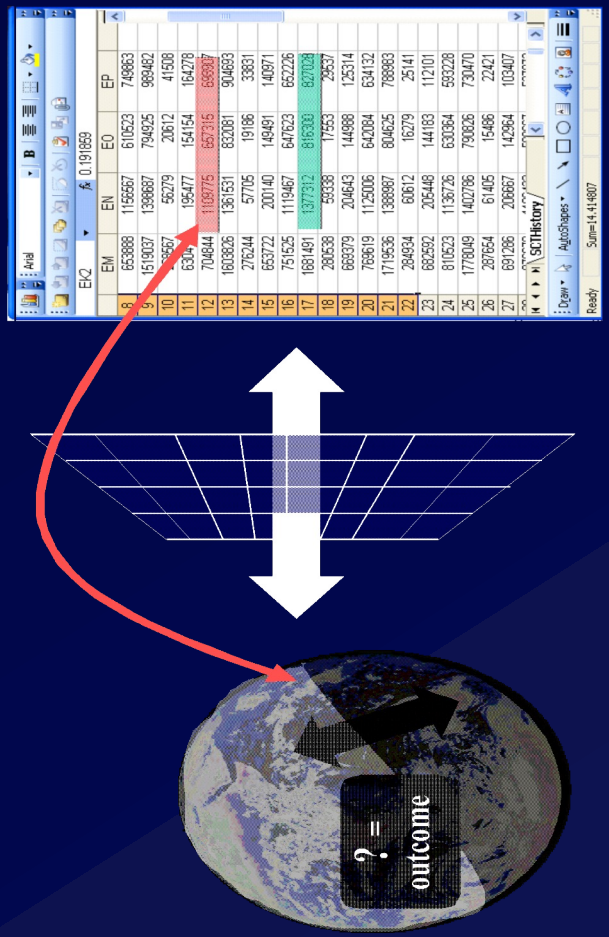


	EM	EN	EO	EP
8	653888	1156567	610523	749863
9	1519037	1398687	794925	989482
10	[illegible]	56279	20612	41508
11	6304[illegible]	195477	154154	164278
12	704844	1109775	657315	698907
13	1603826	1361531	832081	904693
14	276244	57705	19186	33831
15	653722	200140	149491	140971
16	751525	1119467	647623	652226
17	1681491	1377312	816300	827028
18	2805[illegible]	59338	17553	29537
19	66[illegible]379	204643	144988	125314
20	[illegible]69619	1125006	642084	634132
21	1719536	1388987	804625	788983
[illegible]	284934	60612	16279	25141
23	682592	205448	144183	112101
24	810523	1136726	630364	593228
25	1778049	1402786	790826	730470
26	287654	61405	15486	22421
27	691286	206667	142964	103407

Referents

References organized in a *data collection*

… even when reality changes



- Are differences in data about the same entities in reality at different points in time due to:
  - changes in first-order reality ?
  - changes in our understanding of reality ?
  - inaccurate observations ?
  - registration mistakes ?

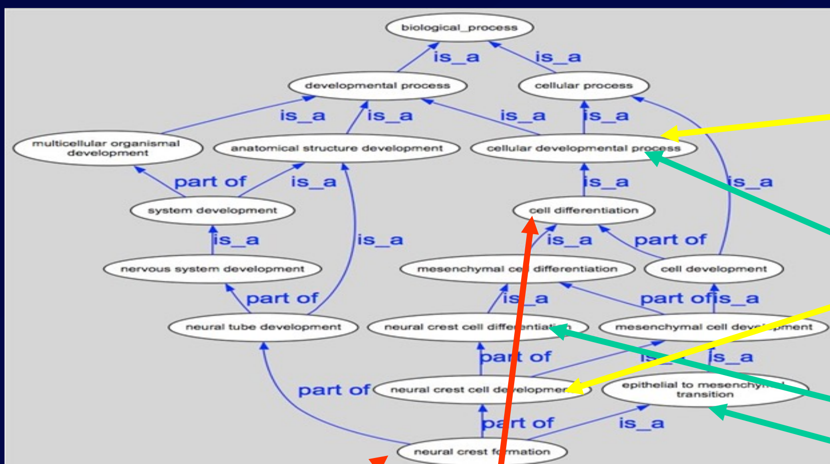
Ceusters W, Smith B. A Realism-Based Approach to the Evolution of Biomedical Ontologies. Proceedings of AMIA 2006, Washington DC, 2006;:121-125.
http://www.referent-tracking.com/RTU/sendfile/?file=CeustersAMIA2006FINAL.pdf

Methods to achieve faithfulness and data clarity

Sources	Data generation	Data organization
Data collection sheets	✓	
Instruction manuals	✓	
Interpretation criteria	✓	
Diagnostic criteria	✓	
Assessment instruments	✓	
Terminologies	✓	✓
Data validation procedures		✓
Data dictionaries		✓
Ontologies	✓	✓

If not used for data collection and organization, these sources can be used post hoc to document, and perhaps increase, the level of data clarity and faithfulness in and comparability of existing data collections.

Ontologies to make data collections comparable



Cases	Characteristics						
	ch1	ch2	ch3	ch4	ch5	ch6	...
case1							
case2							
case3							
case4							
case5							
case6							
...							

Cases	Characteristics						
	ch1	ch2	ch3	ch4	ch5	ch6	...
case1							
case2							
case3							
case4							
case5							
case6							
...							

Cases	Characteristics						
	ch1	ch2	ch3	ch4	ch5	ch6	...
case1							
case2							
case3							
case4							
case5							
case6							
...							

Linking the variables of distinct data collections to a realism-based ontology.

OPMQoL: an Ontology for pain-related disablement, mental health and quality of life

Acknowledgement

The work described is funded in part by grant 1R01DE021917-01A1 from the National Institute of Dental and Craniofacial Research (NIDCR). The content of this presentation is solely the responsibility of the author and does not necessarily represent the official views of the NIDCR or the National Institutes of Health.

Collaborators

Werner Ceusters – Richard Ohrbach
UB (PIs)

Mike T. John – Eric L. Schiffman
University of Minnesota

Vishar Aggarwal
Manchester, UK

Joanna Zakrzewska
London, UK

Thomas List
Malmö, Sweden

Rafael Benoliel
Hadassah, Israel

Background (1)

- July 2008, Toronto:
  - the International RDC/TMD Consortium Network identified a need to incorporate the RDC/TMD diagnostic taxonomy into a comprehensive orofacial pain taxonomy.
- April, 2009, Miami:
  - 'The International Consensus Workshop: Convergence on an Orofacial Pain Taxonomy' participants decided that an adequate treatment of the *ontology* of pain in general, and orofacial pain in particular, together with an appropriate *terminology*, is mandatory to advance the state of the art in diagnosis, treatment and prevention.

Ohrbach R, List T, Goulet J, Svensson P. Recommendations from the International Consensus Workshop: Convergence on an Orofacial Pain Taxonomy. Journal of Oral Rehabilitation. 2010.

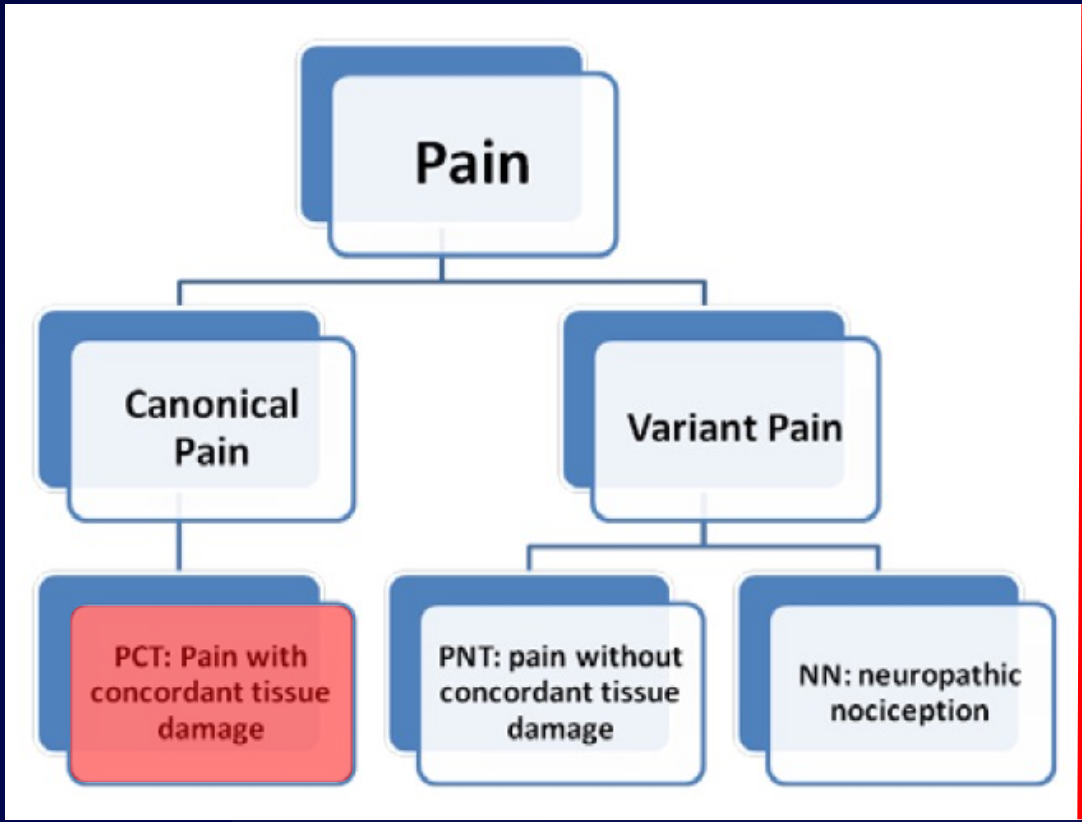
Background (2)

- The following consecutive steps were proposed:
  1. study the terminology and ontology of pain as currently defined,
  2. find ways to make individual data collections more useful for international research,
  3. develop an ontology for integrating knowledge and data over all the known basic and clinical science domains concerning TMD and its relationship to complex disorders, and
  4. expand this ontology to cover all pain-related disorders.

Study the terminology of pain as currently defined

- Starting point - the IASP definition for 'pain':
  - '*an unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage*';
- what asserts:
  - a common phenomenology ('unpleasant sensory and emotional experience') to all instances of pain,
  - the recognition of three distinct subtypes of pain involving, respectively:
    1. actual tissue damage,
    2. what is called 'potential tissue damage', and
    3. a description involving reference to tissue damage whether or not there is such damage.

Results of the ontological analysis



For ontological definitions of these types, see:

Smith B, Ceusters W, Goldberg LJ, Ohrbach R. Towards an Ontology of Pain. In: Mitsu Okada (ed.), Proceedings of the Conference on Logic and Ontology, Tokyo: Keio University Press, February 2011:23-32.
http://www.referent-tracking.com/RTU/sendfile/?file=painTokyo1_27_2011.pdf
http://www.referent-tracking.com/RTU/sendfile/?file=201201PainProject.ppt

Main topics for this lecture

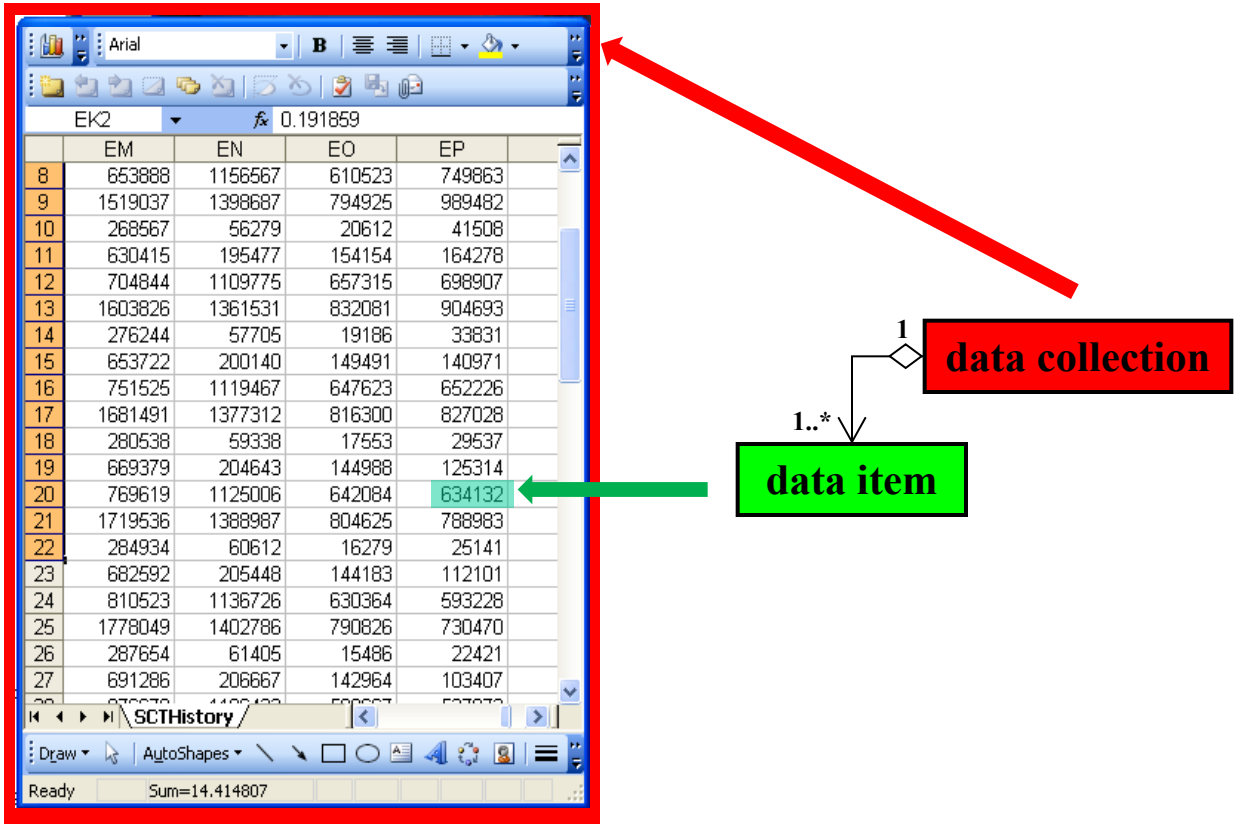
- The following consecutive steps were proposed:
  1. study the terminology and ontology of pain as currently defined,
  2. find ways to make individual data collections more useful for international research,
  3. develop an ontology for integrating knowledge and data over all the known basic and clinical science domains concerning TMD and its relationship to complex disorders, and
  4. expand this ontology to cover all pain-related disorders.

Find ways to make individual data collections more useful for international research

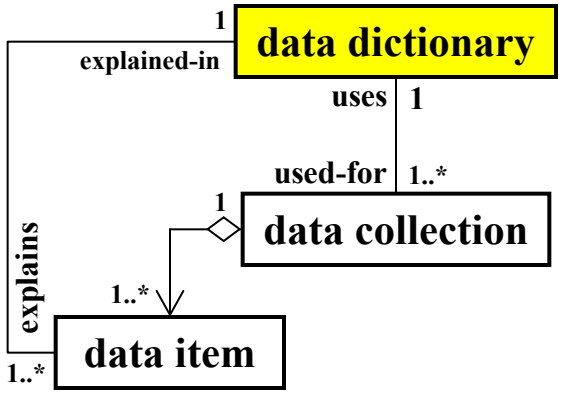
Methodology

- For data collections still to be generated:
  - express all study criteria, variables and possible values in terms of realism-based ontologies
- For existing data collections:
  - analyze data dictionaries, assessment instruments, study criteria and corresponding terminologies,
  - build realism-based application ontologies to link these sources to realism-based reference ontologies.

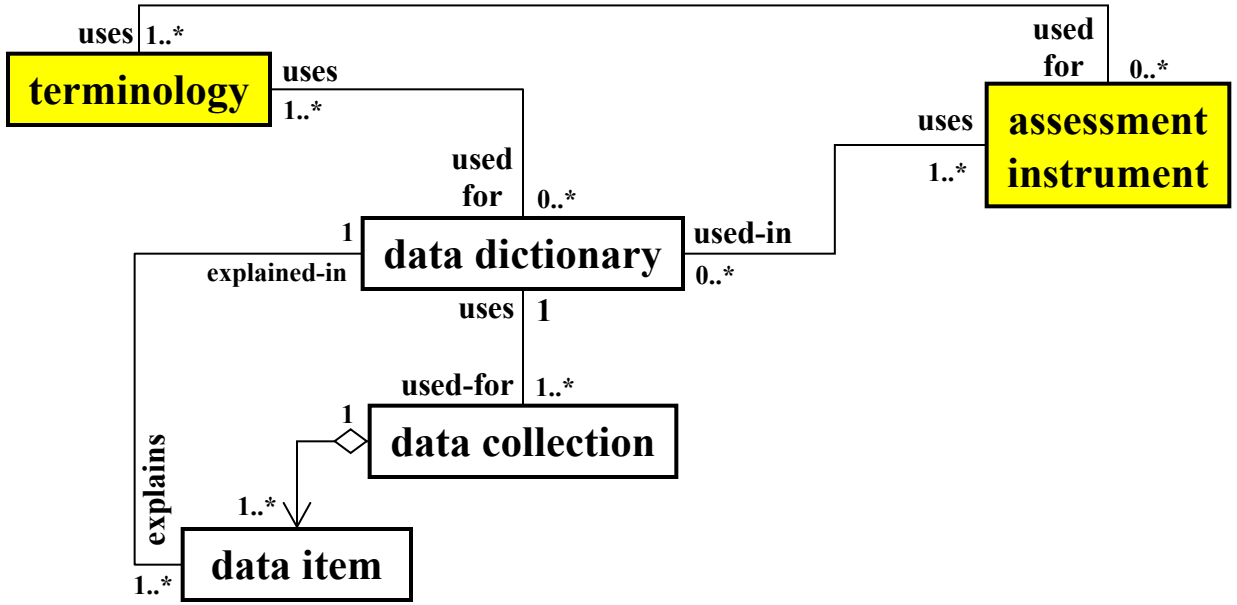
A data collection consists of at least 1 data item, each data item belonging to exactly 1 collection



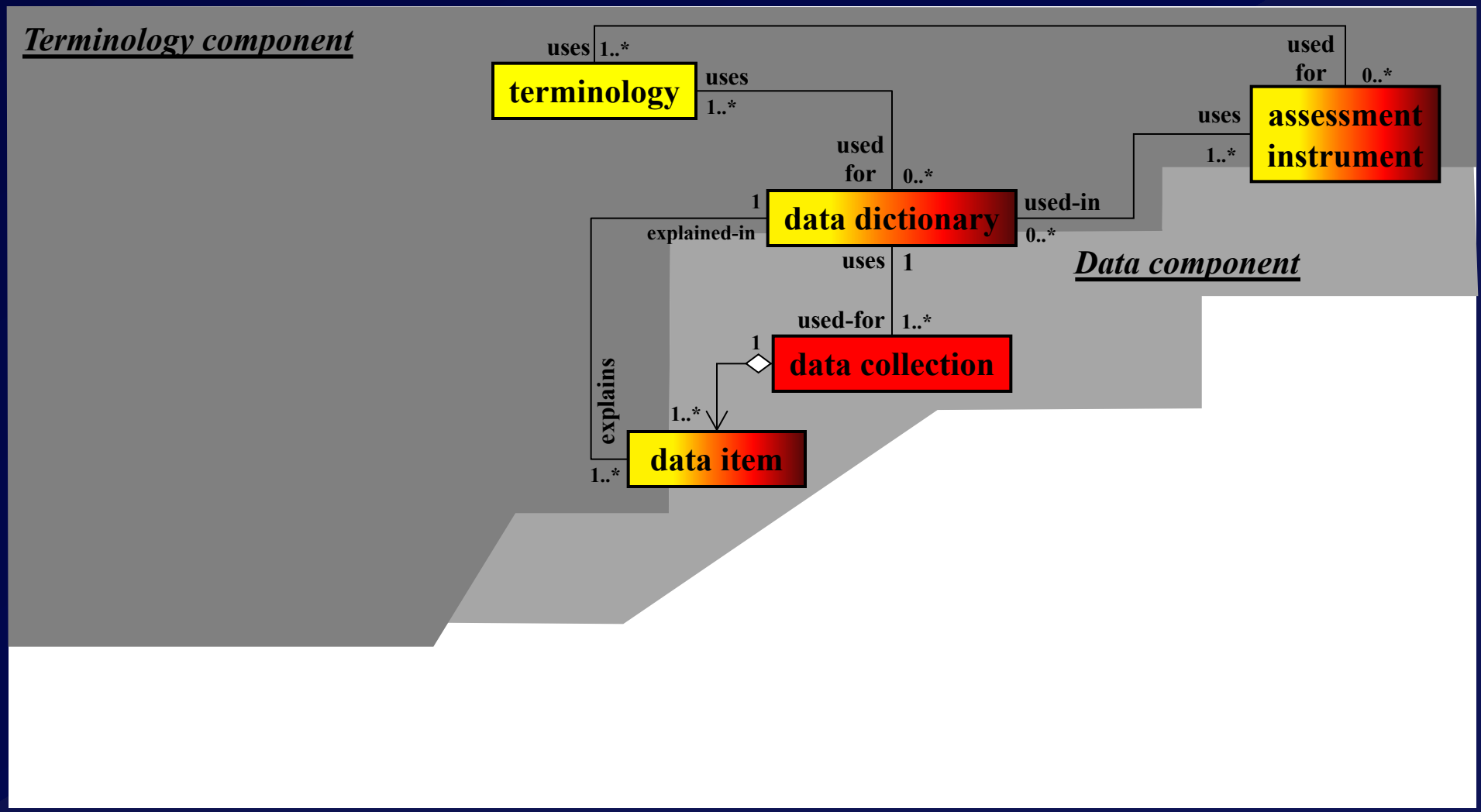
Data dictionaries provide information about data items and data collections



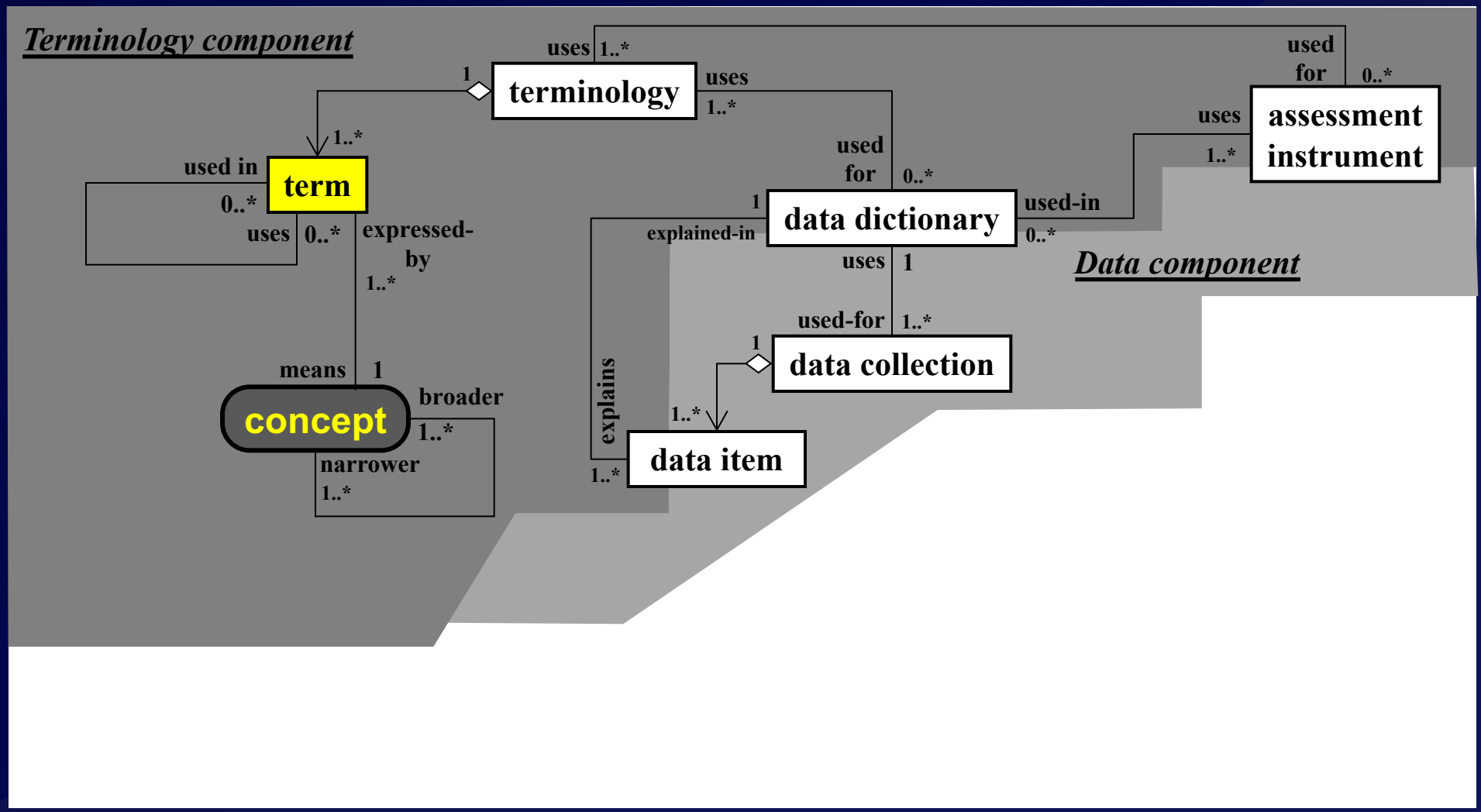
Data dictionaries provide also information about terminologies and assessment instruments used for data generation, in addition to information about the collection's structure



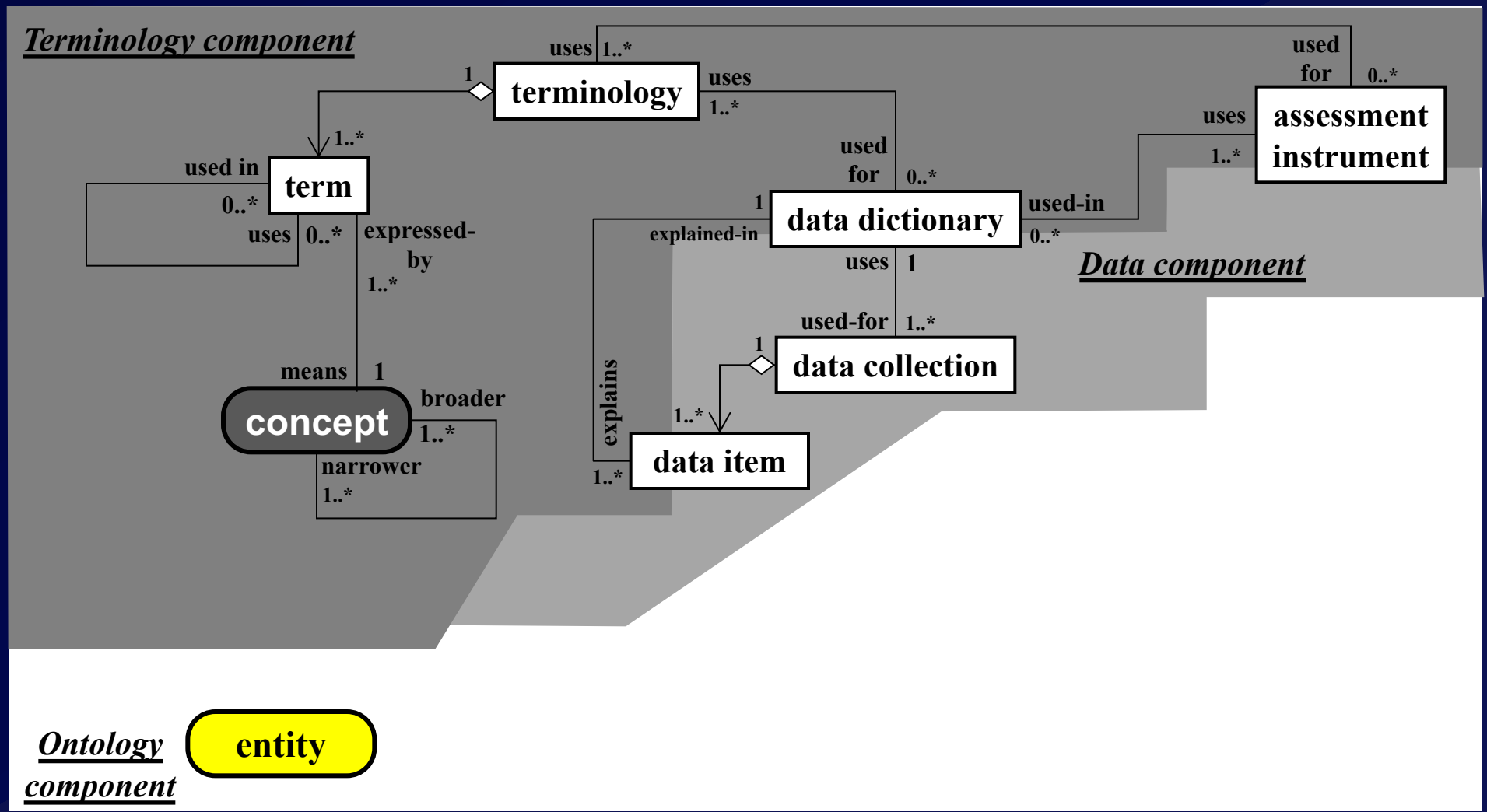
Relation of Terminology component to Data component



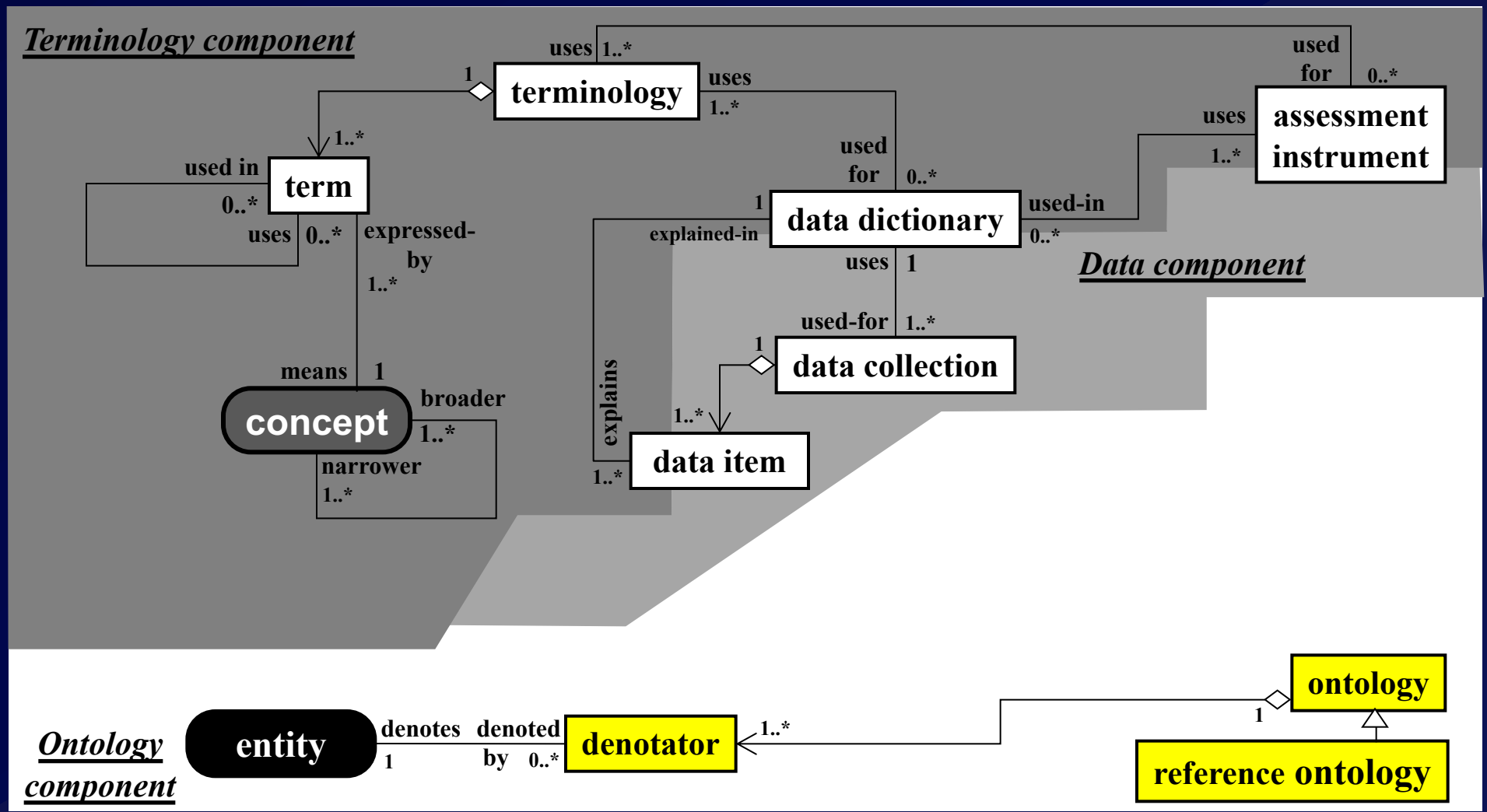
Terminology links terms to '*concepts*'



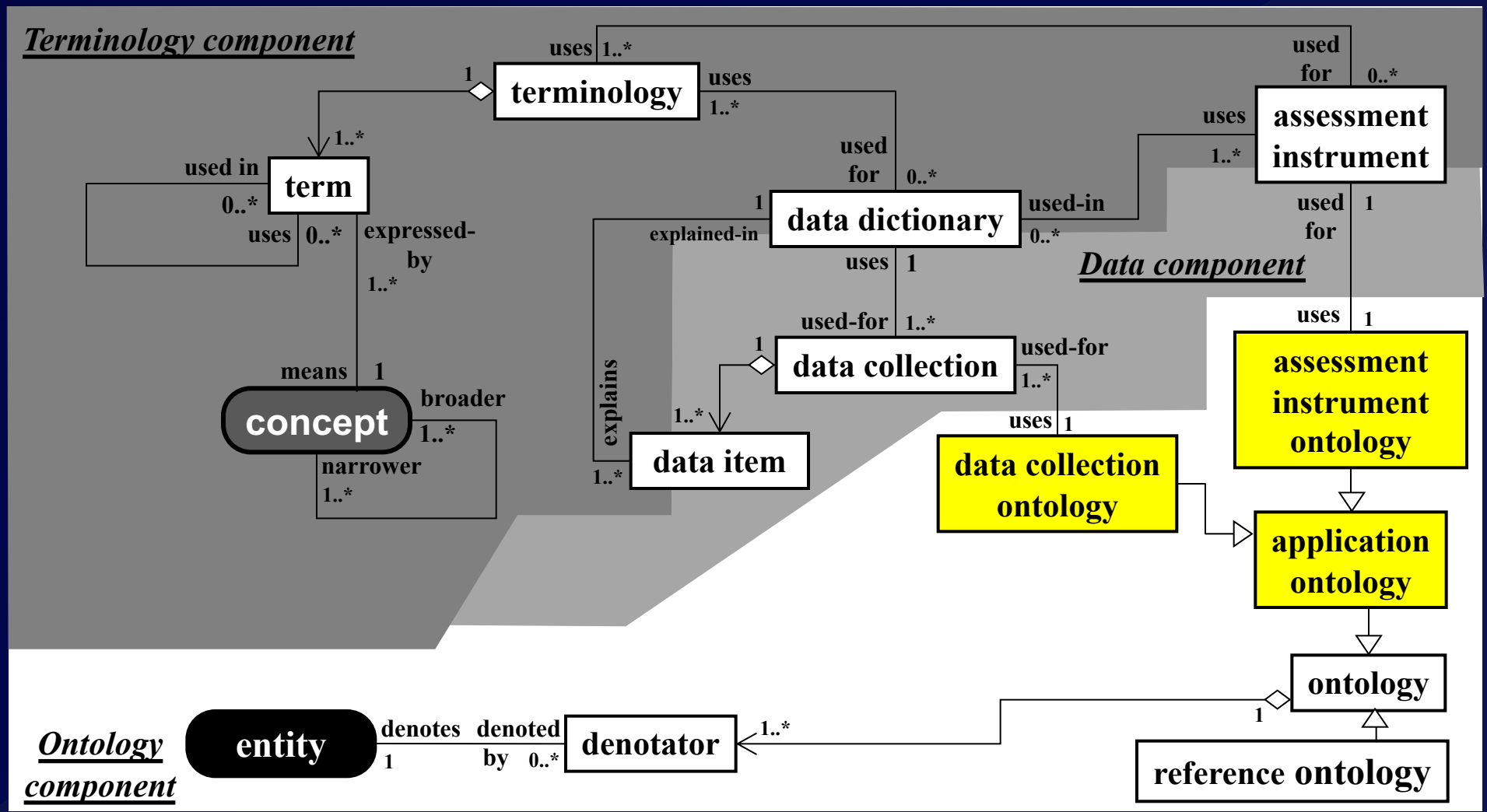
Not 'concepts' are of interest, but entities in reality



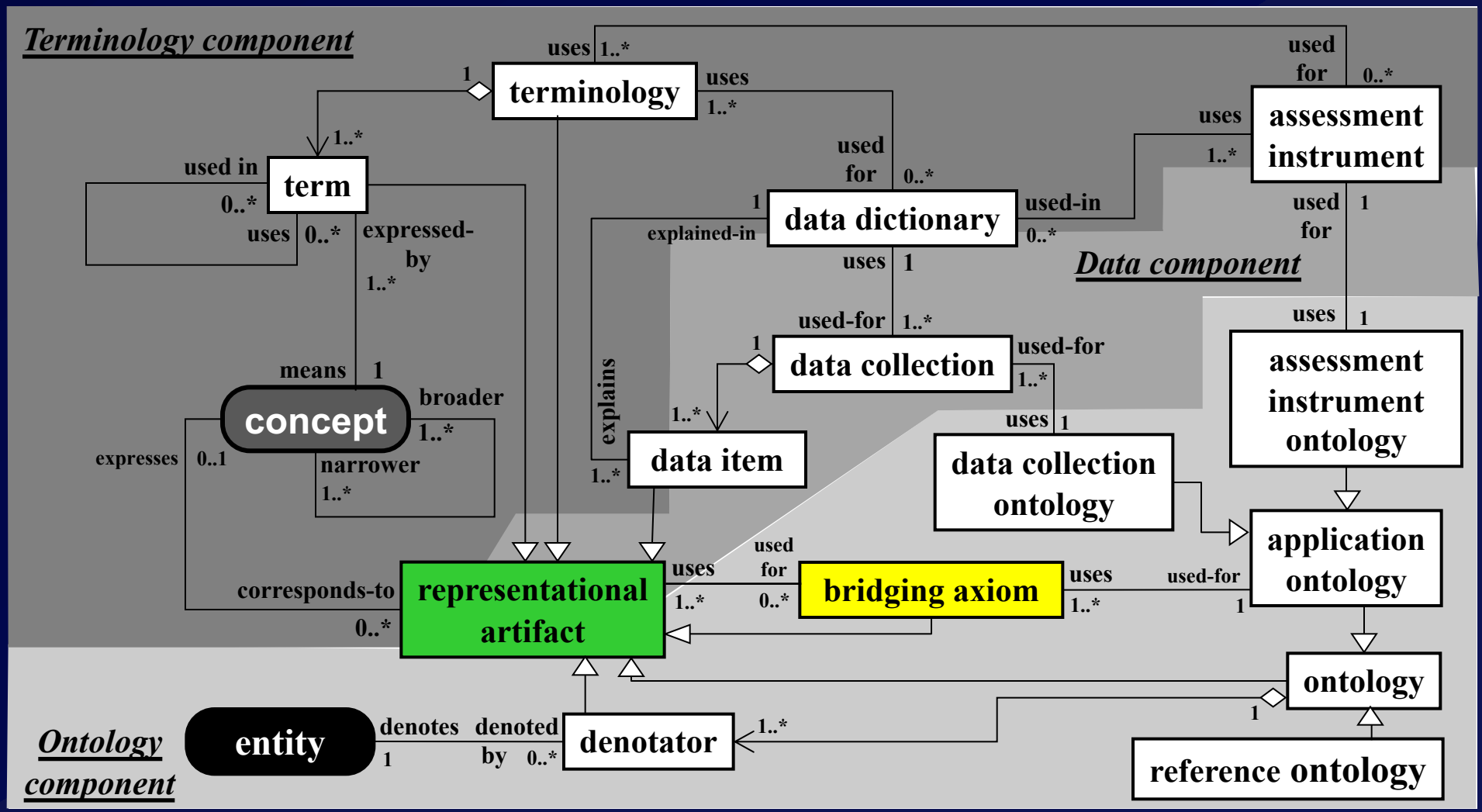
It is real entities that should be denoted in ontologies



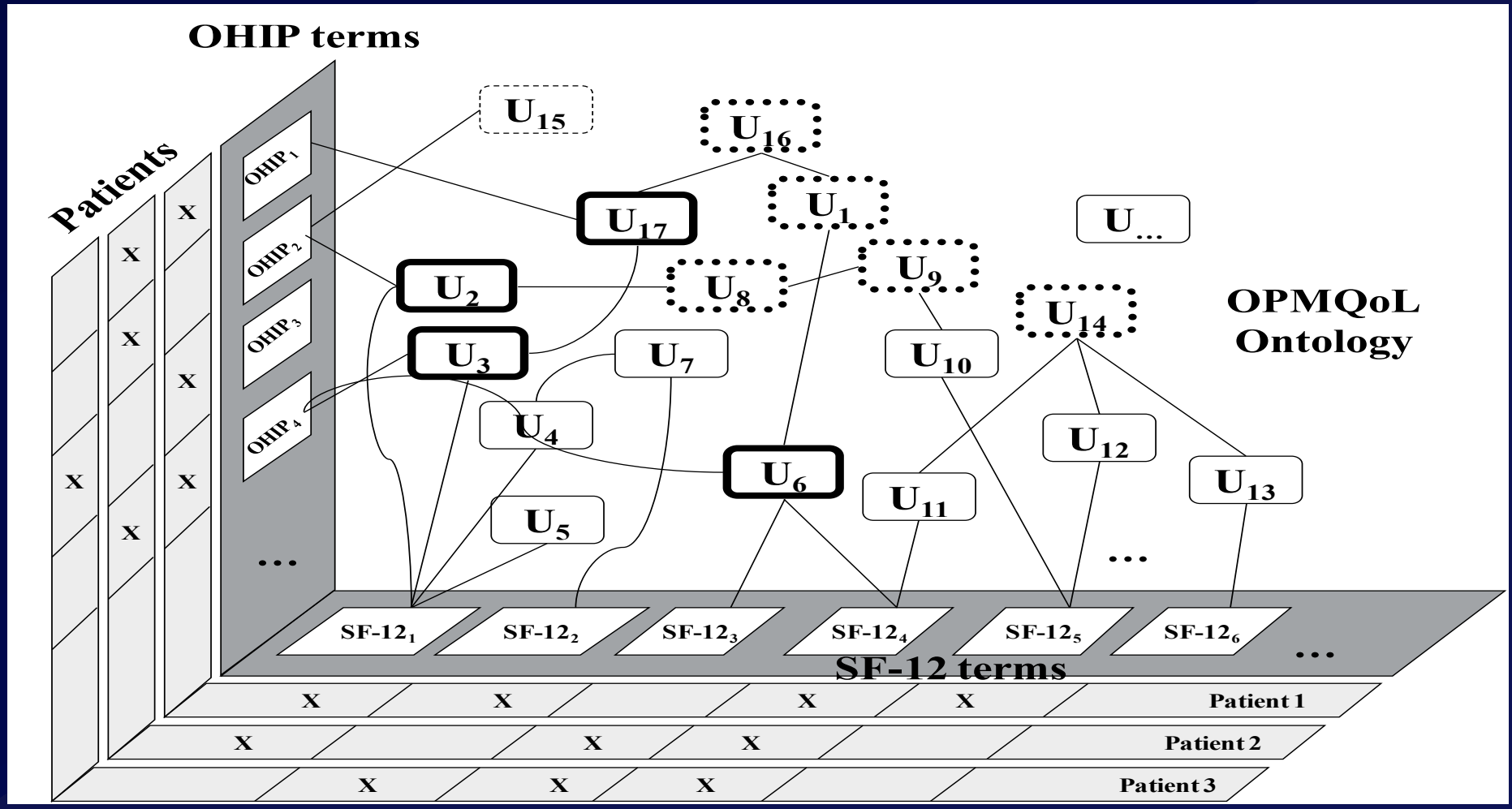
Application ontologies cover the domains of the sources



Bridging axioms link data to ontologies and terminologies

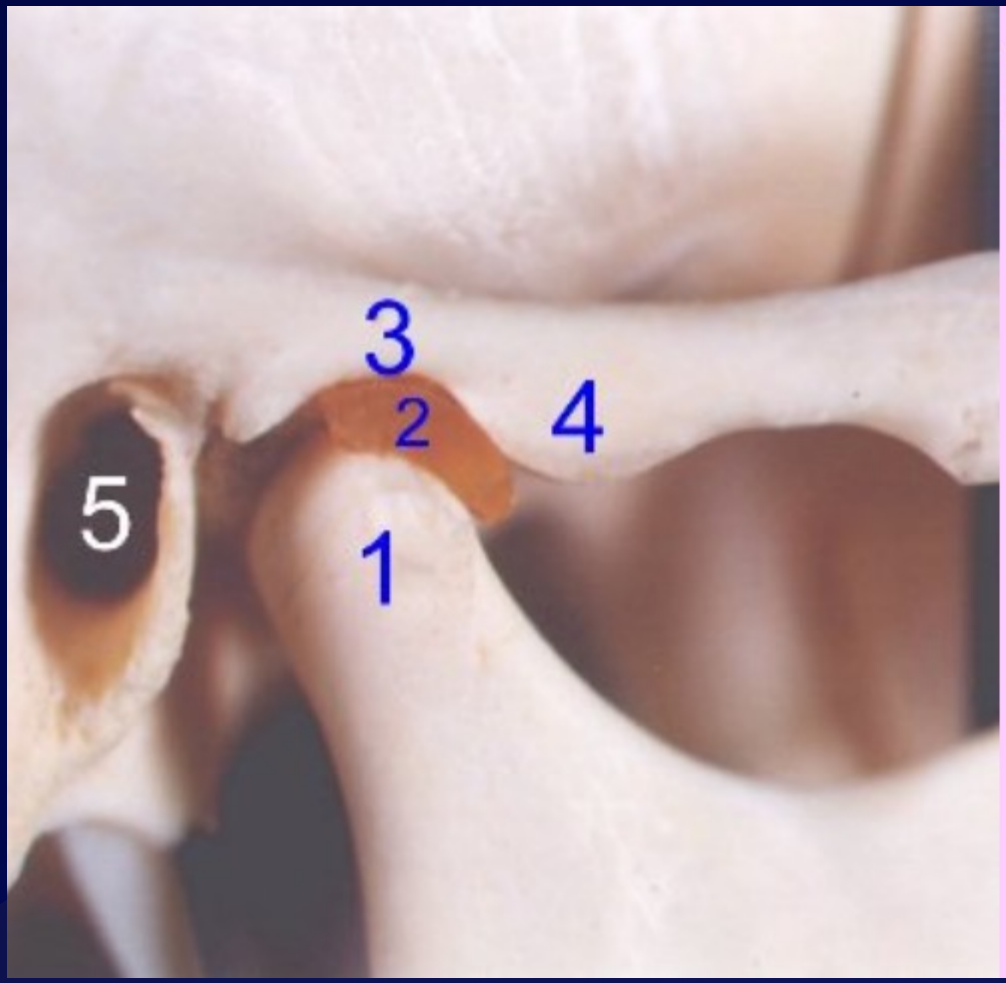


Linking data collections using distinct assessment instruments



Detailed example of the ontology building procedure

Example: assessing TMJ Anatomy

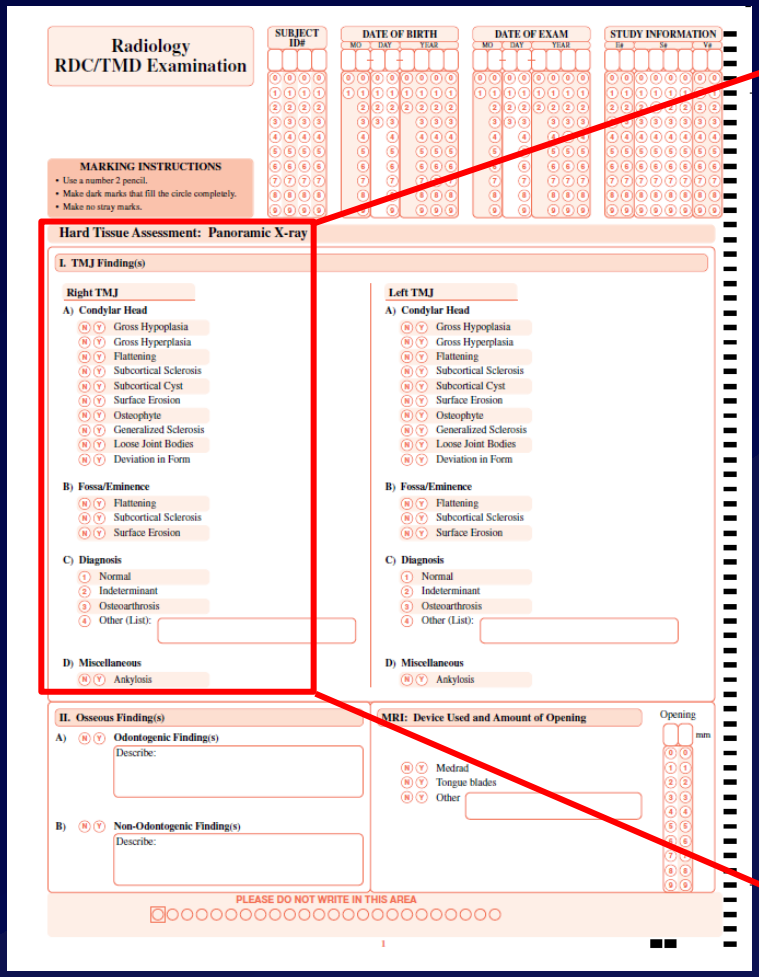


Temporomandibular Joint (TMJ)

1. Mandibular condyle
2. Articular disk
3. Superior joint cavity
4. Articular eminence
5. External ear

Panoramic X-ray of mouth

Radiology RDC/TMD Examination: data collection sheet



Radiology
RDC/TMD Examination

SUBJECT ID# | DATE OF BIRTH (MO DAY YEAR) | DATE OF EXAM (MO DAY YEAR) | STUDY INFORMATION (Tx Tx Vr)

MARKING INSTRUCTIONS
- Use a number 2 pencil.
- Make dark marks that fill the circle completely.
- Make no stray marks.

Hard Tissue Assessment: Panoramic X-ray

I. TMJ Finding(s)

Right TMJ	Left TMJ
A) Condylar Head	A) Condylar Head
(N) (Y) Gross Hypoplasia	(N) (Y) Gross Hypoplasia
(N) (Y) Gross Hyperplasia	(N) (Y) Gross Hyperplasia
(N) (Y) Flattening	(N) (Y) Flattening
(N) (Y) Subcortical Sclerosis	(N) (Y) Subcortical Sclerosis
(N) (Y) Subcortical Cyst	(N) (Y) Subcortical Cyst
(N) (Y) Surface Erosion	(N) (Y) Surface Erosion
(N) (Y) Osteophyte	(N) (Y) Osteophyte
(N) (Y) Generalized Sclerosis	(N) (Y) Generalized Sclerosis
(N) (Y) Loose Joint Bodies	(N) (Y) Loose Joint Bodies
(N) (Y) Deviation in Form	(N) (Y) Deviation in Form
B) Fossa/Eminence	B) Fossa/Eminence
(N) (Y) Flattening	(N) (Y) Flattening
(N) (Y) Subcortical Sclerosis	(N) (Y) Subcortical Sclerosis
(N) (Y) Surface Erosion	(N) (Y) Surface Erosion
C) Diagnosis	C) Diagnosis
(1) Normal	(1) Normal
(2) Indeterminant	(2) Indeterminant
(3) Osteoarthrosis	(3) Osteoarthrosis
(4) Other (List):	(4) Other (List):
D) Miscellaneous	D) Miscellaneous
(N) (Y) Ankylosis	(N) (Y) Ankylosis

II. Osseous Finding(s)

A) (N) (Y) Odontogenic Finding(s)
Describe:

B) (N) (Y) Non-Odontogenic Finding(s)
Describe:

MRI: Device Used and Amount of Opening

(N) (Y) Medrad
(N) (Y) Tongue blades
(N) (Y) Other

Opening mm

PLEASE DO NOT WRITE IN THIS AREA

Hard Tissue Assessment: Panoramic X-ray

I. TMJ Finding(s)

Right TMJ

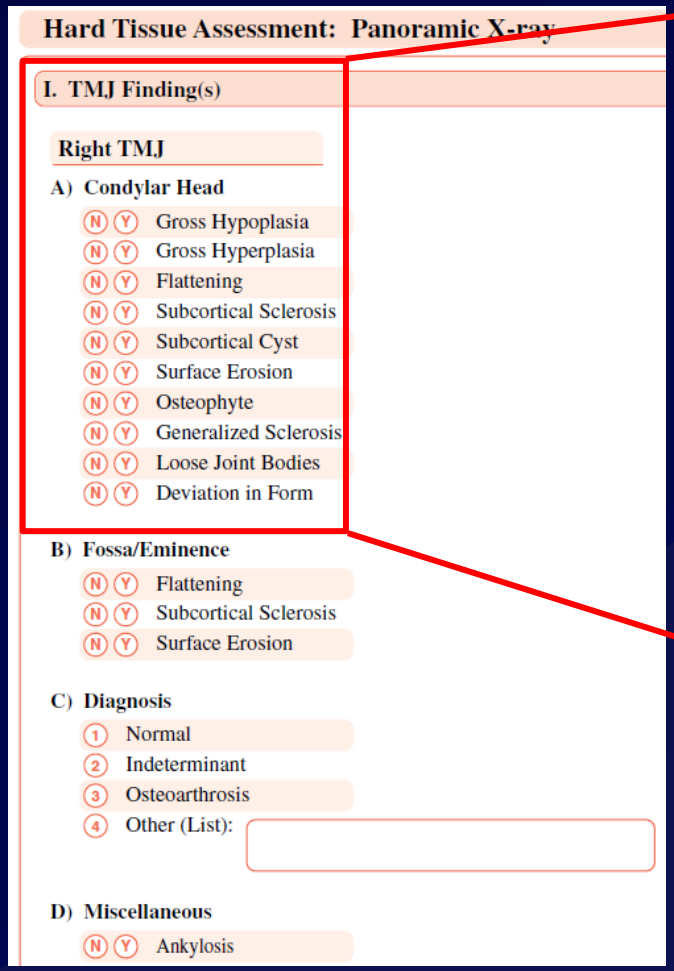
A) Condylar Head
- (N) (Y) Gross Hypoplasia
- (N) (Y) Gross Hyperplasia
- (N) (Y) Flattening
- (N) (Y) Subcortical Sclerosis
- (N) (Y) Subcortical Cyst
- (N) (Y) Surface Erosion
- (N) (Y) Osteophyte
- (N) (Y) Generalized Sclerosis
- (N) (Y) Loose Joint Bodies
- (N) (Y) Deviation in Form

B) Fossa/Eminence
- (N) (Y) Flattening
- (N) (Y) Subcortical Sclerosis
- (N) (Y) Surface Erosion

C) Diagnosis
- (1) Normal
- (2) Indeterminant
- (3) Osteoarthrosis
- (4) Other (List):

D) Miscellaneous
- (N) (Y) Ankylosis

RDC/TMD: a collaborator's data dictionary



I. TMJ Finding(s)

Right TMJ

A.

PanRtIA1	
PanRtIA2	
PanRtIA3	
PanRtIA4	
PanRtIA5	
PanRtIA6	
PanRtIA7	
PanRtIA8	
PanRtIA9	
PanRtIA10	

0 - No
1 - Yes
9 - missing

Gross Hypoplasia
Gross Hyperplasia
Flattening
Subcortical Sclerosis
Sybcortical Cyst
Surface Erosion
Osteophyte
Generalized Sclerosis
Loose Joint Bodies
Deviation in Form

Fieldnames in Eric's data collection

Allowed values for the fields

Anybody sees something disturbing ?

Hard Tissue Assessment: Panoramic X-ray

I. TMJ Finding(s)

Right TMJ

A) Condylar Head

- (N) (Y) Gross Hypoplasia
- (N) (Y) Gross Hyperplasia
- (N) (Y) Flattening
- (N) (Y) Subcortical Sclerosis
- (N) (Y) Subcortical Cyst
- (N) (Y) Surface Erosion
- (N) (Y) Osteophyte
- (N) (Y) Generalized Sclerosis
- (N) (Y) Loose Joint Bodies
- (N) (Y) Deviation in Form

B) Fossa/Eminence

- (N) (Y) Flattening
- (N) (Y) Subcortical Sclerosis
- (N) (Y) Surface Erosion

C) Diagnosis

- (1) Normal
- (2) Indeterminant
- (3) Osteoarthrosis
- (4) Other (List):

D) Miscellaneous

- (N) (Y) Ankylosis

I. TMJ Finding(s)

A. Right TMJ

0 - No
1 - Yes
9 - missing

PanRtIA1		Gross Hypoplasia
PanRtIA2		Gross Hyperplasia
PanRtIA3		Flattening
PanRtIA4		Subcortical Sclerosis
PanRtIA5		Sybcortical Cyst
PanRtIA6		Surface Erosion
PanRtIA7		Osteophyte
PanRtIA8		Generalized Sclerosis
PanRtIA9		Loose Joint Bodies
PanRtIA10		Deviation in Form

This data dictionary alone is not reliable!

Hard Tissue Assessment: Panoramic X-ray

I. TMJ Finding(s)

Right TMJ

A) Condylar Head

- (N) (Y) Gross Hypoplasia
- (N) (Y) Gross Hyperplasia
- (N) (Y) Flattening
- (N) (Y) Subcortical Sclerosis
- (N) (Y) Subcortical Cyst
- (N) (Y) Surface Erosion
- (N) (Y) Osteophyte
- (N) (Y) Generalized Sclerosis
- (N) (Y) Loose Joint Bodies
- (N) (Y) Deviation in Form

B) Fossa/Eminence

- (N) (Y) Flattening
- (N) (Y) Subcortical Sclerosis
- (N) (Y) Surface Erosion

C) Diagnosis

- (1) Normal
- (2) Indeterminant
- (3) Osteoarthrosis
- (4) Other (List):

D) Miscellaneous

- (N) (Y) Ankylosis

I. TMJ Finding(s)

A. Right TMJ

PanRtIA1	Gross Hypoplasia
PanRtIA2	Gross Hyperplasia
PanRtIA3	Flattening
PanRtIA4	Subcortical Sclerosis
PanRtIA5	Sybcortical Cyst
PanRtIA6	Surface Erosion
PanRtIA7	Osteophyte
PanRtIA8	Generalized Sclerosis
PanRtIA9	Loose Joint Bodies
PanRtIA10	Deviation in Form

0 - No
1 - Yes
9 - missing

That these variables are about the ***condylar head*** of the TMJ is 'lost in translation'!

‘meaning’ of values in data collections

I. TMJ Finding(s)

A.

	Right TMJ	
PanRtIA1		Gross Hypoplasia
PanRtIA2		Gross Hyperplasia
PanRtIA3		Flattening
PanRtIA4		Subcortical Sclerosis
PanRtIA5		Sybcortical Cyst
PanRtIA6		Surface Erosion
PanRtIA7		Osteophyte
PanRtIA8		Generalized Sclerosis
PanRtIA9		Loose Joint Bodies
PanRtIA10		Deviation in Form

0 - No
1 - Yes
9 - missing

‘The patient with patient identifier ‘PtID4’ is stated to have had a panoramic X-ray of the mouth which is interpreted to show subcortical sclerosis of that patient’s condylar head of the right temporomandibular joint’

meaning

PainAnalysisExample

Home Insert Page Layout Formulas Data Review View Acrobat

A1 | PtID

	A	B	C	D	E	F	G	H	I	J	K	…
1	PtID	PanRtIA1	PanRtIA2	PanRtIA3	PanRtIA4	PanRtIA5	PanRtIA6	PanRtIA7	PanRtIA8	PanRtIA9	PanRtIA10	
2	PtID1	0	0	0	1	9	0	0	0	0	0	
3	PtID2	0	1	0	1	0	0	0	0	0	0	
4	PtID3	0	0	0	9	0	0	0	0	0	0	
5	PtID4	0	0	0	1	1	0	9	0	1	1	
6	PtID5	0	1	9	1	0	0	0	1	0	0	
7	PtID6	0	0	0	1	2	0	0	0	0	0	
8	PtID7	0	0	0	1	0	0	0	0	0	0	
9	…											

Objectives of the 'sources' analysis

- Find for each value V in the data collections all possible configurations of entities (according to our best scientific understanding) for which the following can be true:
  - V
  - 'it is stated that V'
- Describe these possible configurations by means of sentences from a formal language that mimic the structure of reality.

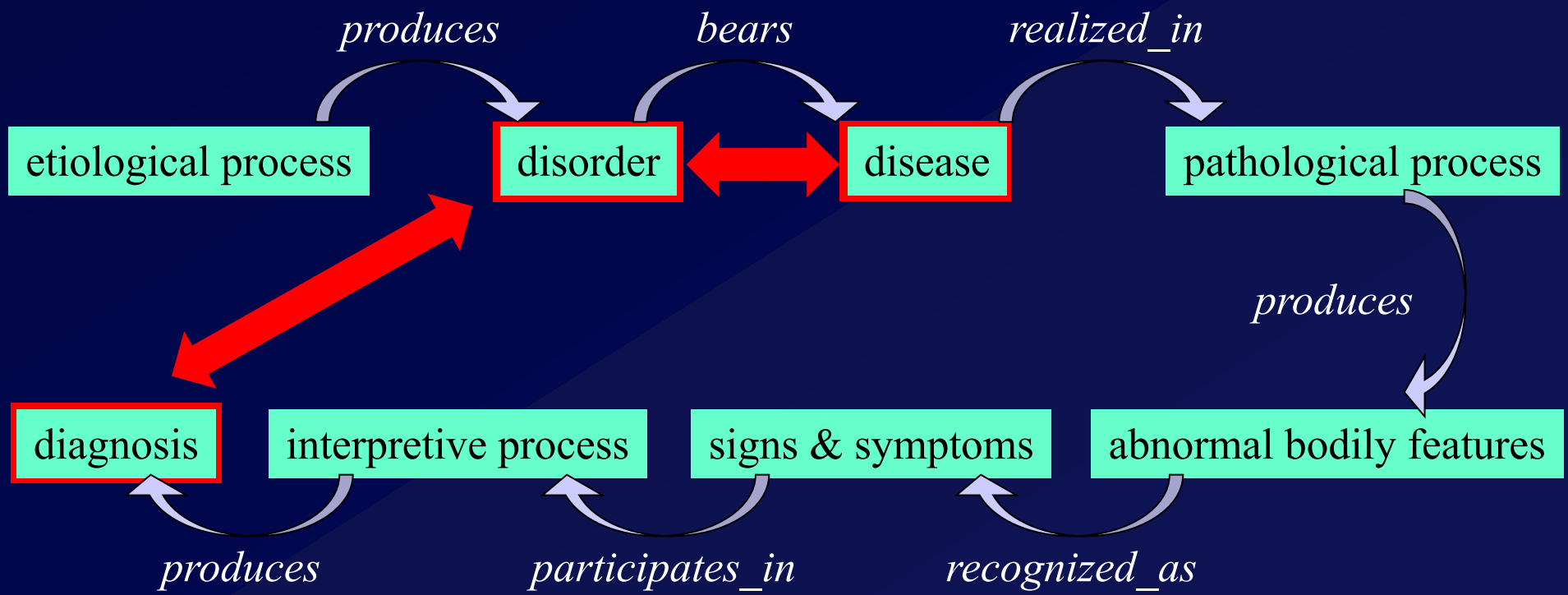
Objectives of the 'sources' analysis (2)

- For example,
  - for the value stating that '*The patient with patient identifier 'PtID4' has had a panoramic X-ray of the mouth which is interpreted to show subcortical sclerosis of that patient's condylar head of the right temporomandibular joint*' to be true,
  - this statement must have been made,
  - for the statement to be true, there must have been that patient, an X-ray, etc, …
  - BUT! It is not necessarily true that that patient has indeed the sclerosis as diagnosed.

Methodology

1. Formulate for each variable in the data collection a sentence explaining as accurately as possible what the variable stands for,
2. list the entities in reality that the terms in the sentence denote,
3. list recursively for all entities listed further entities that ontologically must exist for the entity under scrutiny to exist,
4. classify all entities in terms of realism-based ontologies (RBO),
5. specify all obtaining relationships between these entities,
6. outline all possible configurations of such entities for the sentence to be true.

RBO (1): Ontology of General Medical Science



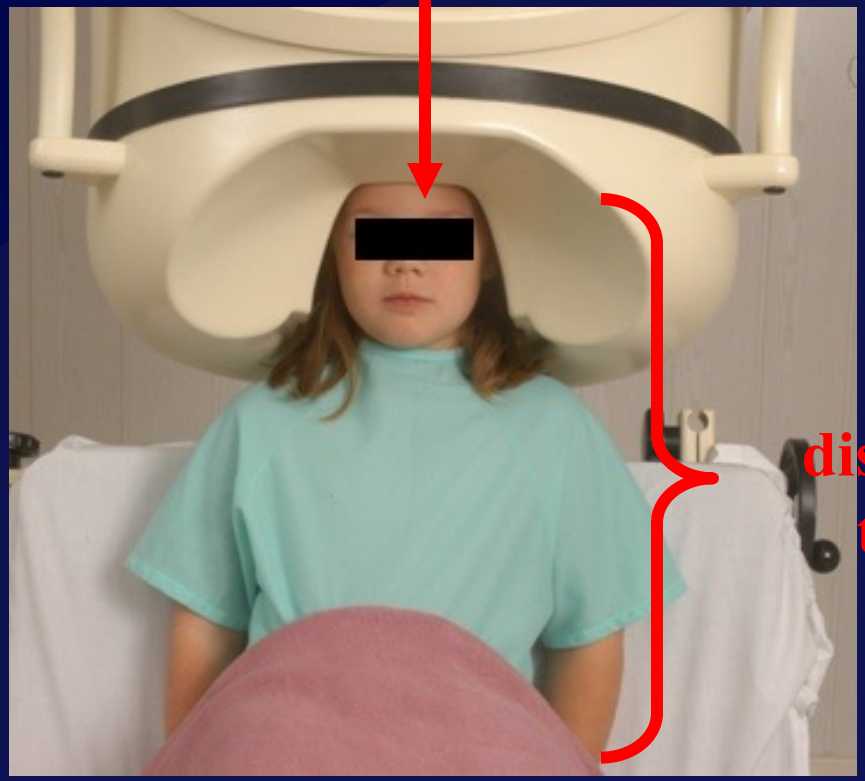
Scheuermann R, Ceusters W, Smith B. Toward an Ontological Treatment of Disease and Diagnosis. 2009 AMIA Summit on Translational Bioinformatics, San Francisco, California, March 15-17, 2009;: 116-120.
http://www.referent-tracking.com/RTU/sendfile/?file=AMIA-0075-T2009.pdf

http://code.google.com/p/ogms/

No conflation of *diagnosis*, *disease*, and *disorder*

The diagnosis is here

The disorder is there



The disease is there

OGMS: Disorder

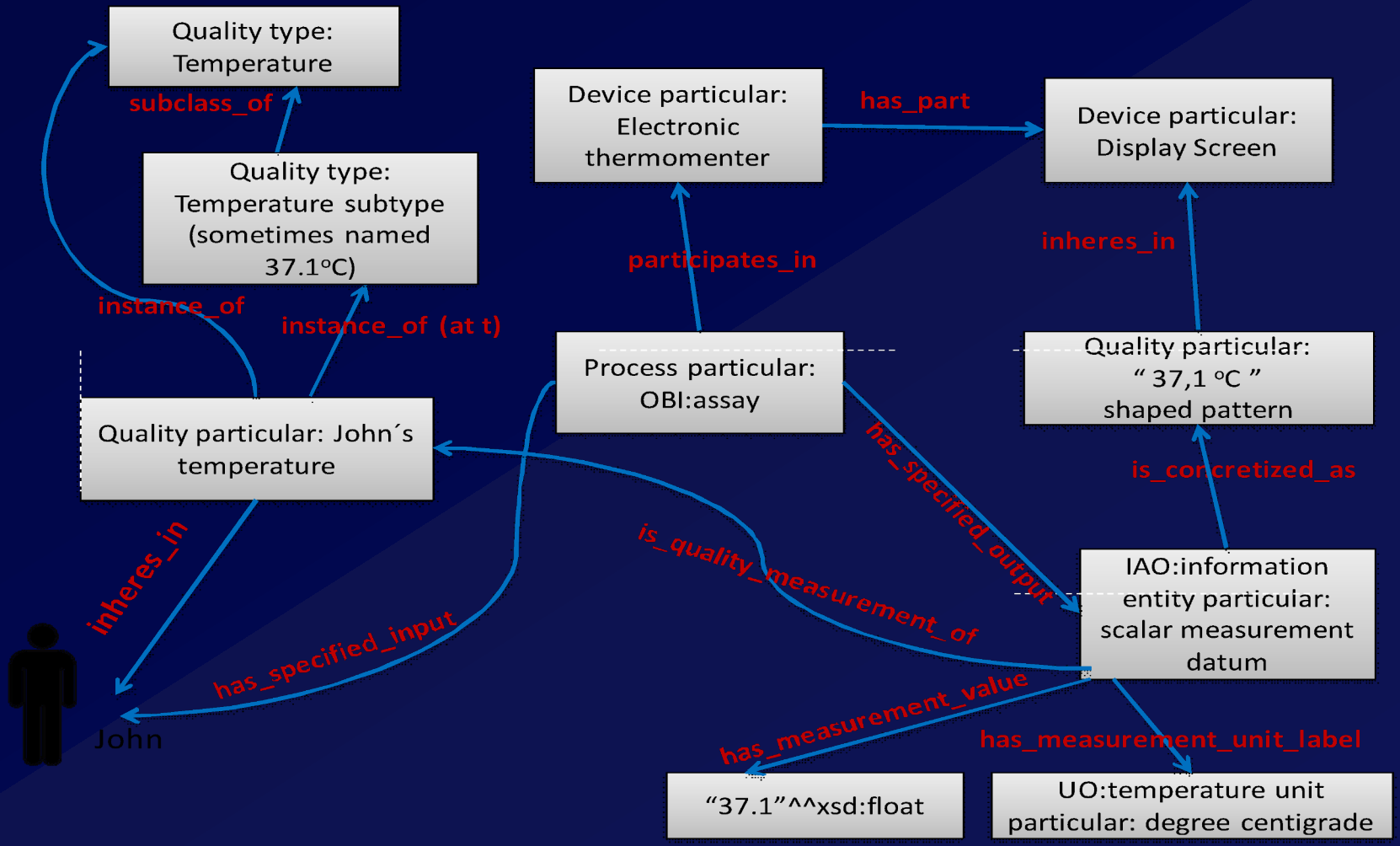
- **Disorder** =def. – A causally linked combination of physical components that is
  - (a) clinically abnormal and
  - (b) maximal, in the sense that it is not a part of some larger such combination.

- - something is clinically abnormal if:
  - (1) is not part of the life plan for an organism of the relevant type (unlike aging or pregnancy),
  - (2) is causally linked to an elevated risk either of pain or other feelings of illness, or of death or dysfunction, and
  - (3) is such that the elevated risk exceeds a certain threshold level.

OGMS: Diagnosis

- **Clinical Picture =def.** – A representation of a clinical phenotype that is inferred from the combination of laboratory, image and clinical findings about a given patient.
- **Diagnosis =def. –**
  - A conclusion of an interpretive process that has as input a clinical picture of a given patient and as output an assertion to the effect that the patient has a disease of such and such a type.

RBO (2): (cleaned up) Ontology of Biomedical Investigations



Step 1: formulate a statement

I. TMJ Finding(s)

A.

	Right TMJ	
PanRtIA1		Gross Hypoplasia
PanRtIA2		Gross Hyperplasia
PanRtIA3		Flattening
PanRtIA4		Subcortical Sclerosis
PanRtIA5		Sybcortical Cyst
PanRtIA6		Surface Erosion
PanRtIA7		Osteophyte
PanRtIA8		Generalized Sclerosis
PanRtIA9		Loose Joint Bodies
PanRtIA10		Deviation in Form

0 - No
1 - Yes
9 - missing

'The patient with patient identifier 'PtID4' is stated to have had a panoramic X-ray of the mouth which is interpreted to show subcortical sclerosis of that patient's condylar head of the right temporomandibular joint'

meaning

PainAnalysisExample

Home Insert Page Layout Formulas Data Review View Acrobat

A1 | PtID

	A	B	C	D	E	F	G	H	I	J	K	
1	PtID	PanRtIA1	PanRtIA2	PanRtIA3	PanRtIA4	PanRtIA5	PanRtIA6	PanRtIA7	PanRtIA8	PanRtIA9	PanRtIA10	...
2	PtID1	0	0	0	1	9	0	0	0	0	0	
3	PtID2	0	1	0	1	0	0	0	0	0	0	
4	PtID3	0	0	0	9	0	0	0	0	0	0	
5	PtID4	0	0	0	1	1	0	9	0	1	1	
6	PtID5	0	1	9	1	0	0	0	1	0	0	
7	PtID6	0	0	0	1	2	0	0	0	0	0	
8	PtID7	0	0	0	1	0	0	0	0	0	0	
9	...											

Step 2 (1): list the entities denoted

- *1(The patient) with 2(patient identifier 'PtID4') 3(is stated) 4(have had) a 5(panoramic X-ray) of 6(the mouth) which 7(is interpreted) to 8(show) 9(subcortical sclerosis of 10(that patient's condylar head of the 11(right temporomandibular joint)))'*

CLASS	INSTANCE IDENTIFIER
person	IUI-1
patient identifier	IUI-2
assertion	IUI-3
technically investigating	IUI-4
panoramic X-ray	IUI-5
mouth	IUI-6
interpreting	IUI-7
seeing	IUI-8
diagnosis	IUI-9
condylar head of right TMJ	IUI-10
right TMJ	IUI-11

notes: colors have no meaning here, just provide easy reference,
this first list can be different, any such differences being resolved in step 3

Step 2 (2): provide *directly referential* descriptions

CLASS	INSTANCE IDENTIFIER	DIRECTLY REFERENTIAL DESCRIPTIONS
person	IUI-1	the person to whom IUI-2 is assigned
patient identifier	IUI-2	the patient identifier of IUI-1
assertion	IUI-3	'the patient with patient identifier PtID4 has had a panoramic X-ray of the mouth which is interpreted to show subcortical sclerosis of that patient's right temporomandibular joint'
technically investigating	IUI-4	the technically investigating of IUI-6
panoramic X-ray	IUI-5	the panoramic X-ray that resulted from IUI-4
mouth	IUI-6	the mouth of IUI-1
interpreting	IUI-7	the interpreting of the signs exhibited by IUI-5
seeing	IUI-8	the seeing of IUI-5 which led to IUI-7
diagnosis	IUI-9	the diagnosis expressed by means of IUI-3
condylar head of right TMJ	IUI-10	the condylar head of the right TMJ of IUI-1
right TMJ	IUI-11	the right TMJ of IUI-1

Step 3: identify further entities that ontologically must exist for each entity under scrutiny to exist.

assigner role	IUI-12	the assigner role played by the entity while it performed IUI-21
assigning	IUI-21	the assigning of IUI-2 to IUI-1 by the entity with role IUI-12
asserting	IUI-20	the asserting of IUI-3 by the entity with asserter role IUI-13
asserter role	IUI-13	the asserter role played by the entity while it performed IUI-20
investigator role	IUI-14	the investigator role played by the entity while it performed IUI-4
panoramic X-ray machine	IUI-15	the panoramic X-ray machine used for performing IUI-4
image bearer	IUI-16	the image bearer in which IUI-5 is concretized and that participated in IUI-8
interpreter role	IUI-17	the interpreter role played by the entity while it performed IUI-7
perceptor role	IUI-18	the perceptor role played by the entity while it performed IUI-8
diagnostic criteria	IUI-19	the diagnostic criteria used by the entity that performed IUI-7 to come to IUI-9
study subject role	IUI-22	the study subject role which inheres in IUI-1

Step 3: some remarks

- interpreter role, perceptor role, …
  - reference to roles rather than the entity in which the roles inhere because it may be the same entity and one should not assign several IUIs to the same entity
- each description follows similar principles as Aristotelian definitions but is about particulars rather than universals

Step 4: classify all entities in terms of realism-based ontologies

CLASS	HIGHER CLASS
person	BFO: Object
patient identifier	IAO: Information Content Entity
assertion	IAO: Information Content Entity
technically investigating	OBI: Assay
panoramic X-ray	IAO: Image
mouth	FMA: Mouth
interpreting	MFO: Assessing
seeing	BFO: Process
diagnosis	IAO: Information Content Entity
condylar head of right TMJ	FMA: Right condylar process of mandible
right TMJ	FMA: Right temporomandibular joint
assigner role	BFO: Role
assigning	BFO: Process
study subject role	OBI: Study subject role

- requires more ontological and philosophical skills than domain expertise or expertise with Protégé,
- not just term matching

Step 5: specify relationships between these entities

- For instance:
  - at least during the taking of the X-ray the study subject role inheres in the patient being investigated:
    - IUI-23 ***inheres-in*** IUI-1 ***during*** t1
  - the patient participates at that time in the investigation
    - IUI-4 ***has-participant*** IUI-1 ***during*** t1
- These relations need to follow the principles of the Relation Ontology.

Smith B, Ceusters W, Klagges B, Koehler J, Kumar A, Lomax J, Mungall C, Neuhaus F, Rector A, Rosse C. Relations in biomedical ontologies, Genome Biology 2005, 6:R46.

Step 6: outline all possible configurations of such entities for the sentence to be true (a one semester course on its own)

- Such outlines are collections of relational expressions of the sort just described,
- Variant configurations for the example:
  - perceptor and interpreter are the same or distinct human beings,
  - the X-ray machine is unreliable and produced artifacts which the interpreter thought to be signs motivating his diagnosis, while the patient has indeed the disorder specified by the diagnosis (the clinician was lucky)
  - …

Conclusion

- Realism-based ontology has a lot to offer to make data collections comparable and unambiguously understandable.
- It is hard !
- How far one needs to go depends on the purposes.
  - ideally: an analysis should be such that it can accommodate ALL purposes, i.e. the analysis should be independent of any purpose;
    - distinction between reference ontologies and application ontologies.