Ph.D. Defense

Terminology and Ontology for Cultural Heritage: Application to Chinese Ceramic Vessels

Candidate: Tong Wei

Supervisors: Christophe Roche, Yangli Jia

Domain: Computer Science

Topics: Ontology, Terminology, Semantic Web, Linked Data, Cultural Heritage, Digital Humanities

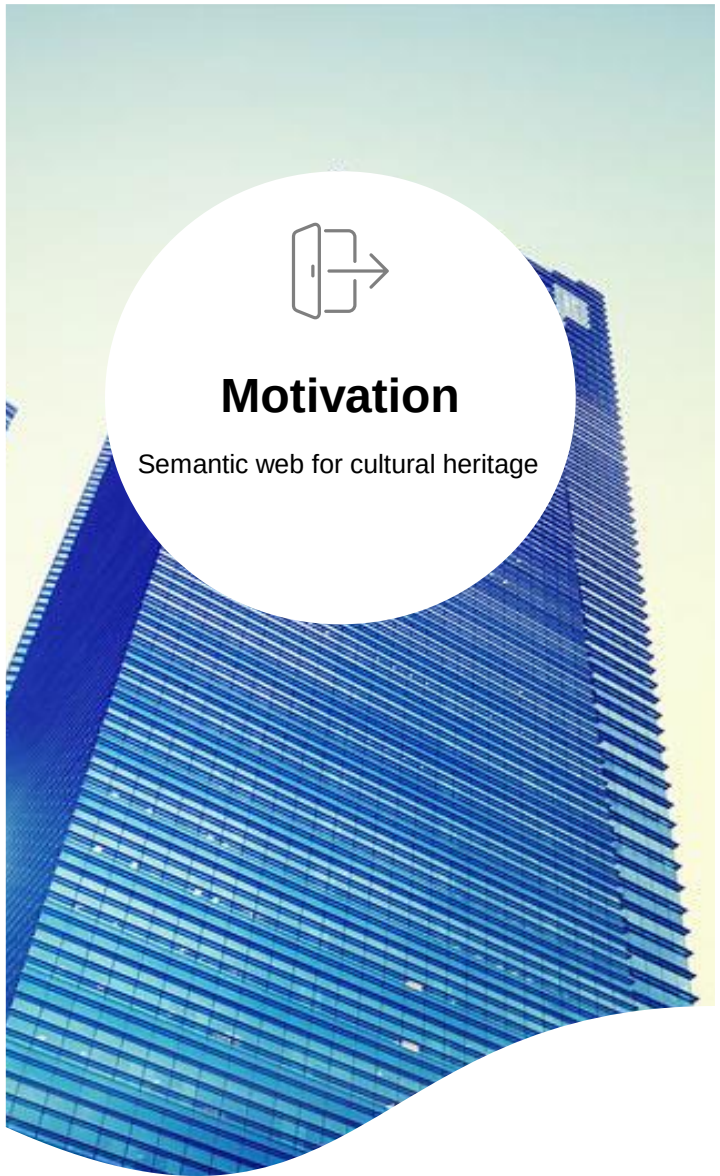
http://www.dh.ketrc.com/

December 04, 2020

CONTENTS

01

Motivation and Research questions



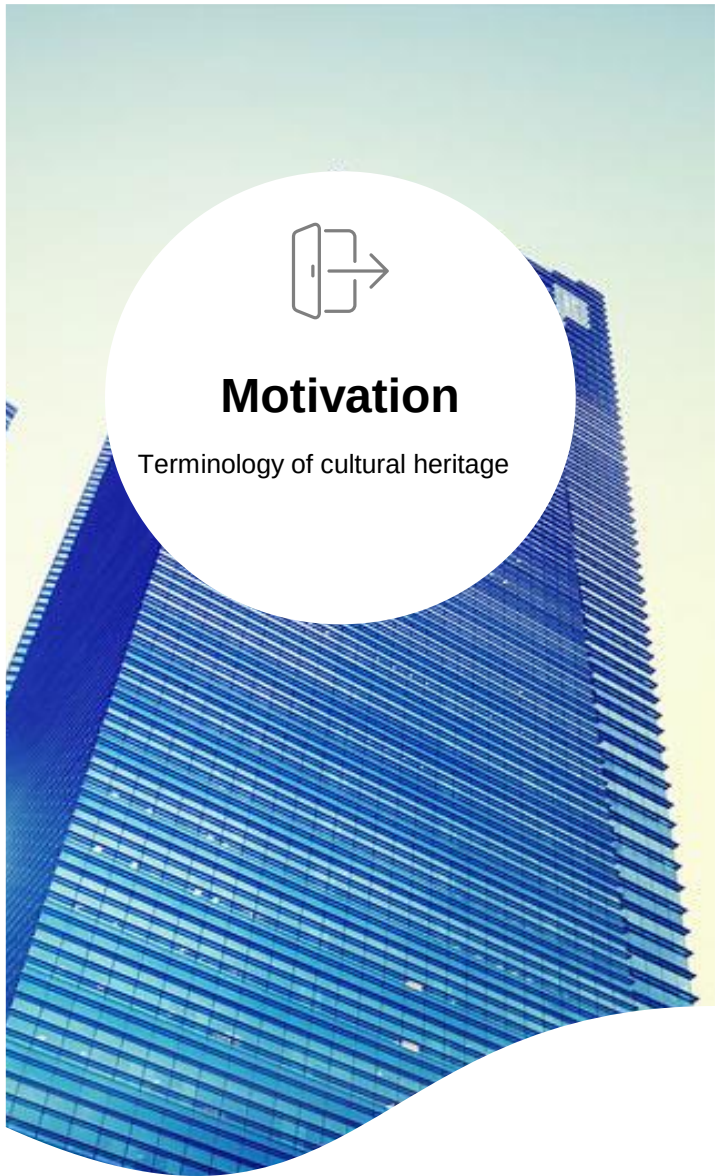
Motivation

Semantic web for cultural heritage

1. Publish open and linked data about the Chinese ceramic vessels of the Ming and Qing Dynasties, as well as the terms denoting them, using the standards of the Semantic Web.

Heritage institutions wish not only to display objects and simple descriptions (drawn from metadata) but also allow for understanding relationships between objects (created by semantically interrelated metadata)

Chinese ceramic vessels are a wealthy domain. Yet it lacks knowledge representation models (ontologies) to capture Chinese pottery concepts, express them in Semantic Web compatible interchange formats, and make them shareable and linkable to other data.



2. Challenge of building knowledge-based terminological resources for communication and knowledge sharing.

For experts or students with distinct language backgrounds, it is difficult to understand the objects denoted by the terms only through the terms.

Naming approach could reflect characteristics of ceramics, but it is not conducive to communication with experts and students of archaeology.

"*dynasty + kiln + glaze + colour + decoration + shape + texture + type*"

Research questions

- What are the theoretical and methodological assumptions underlying the creation of an ontoterminology in the domain of Chinese ceramic vessels?

Research questions

How to create a domain ontology of Chinese ceramic vases following the approach of ontoterminology?

How to take into account the way of thinking of humanists in building terminology and conceptualization?

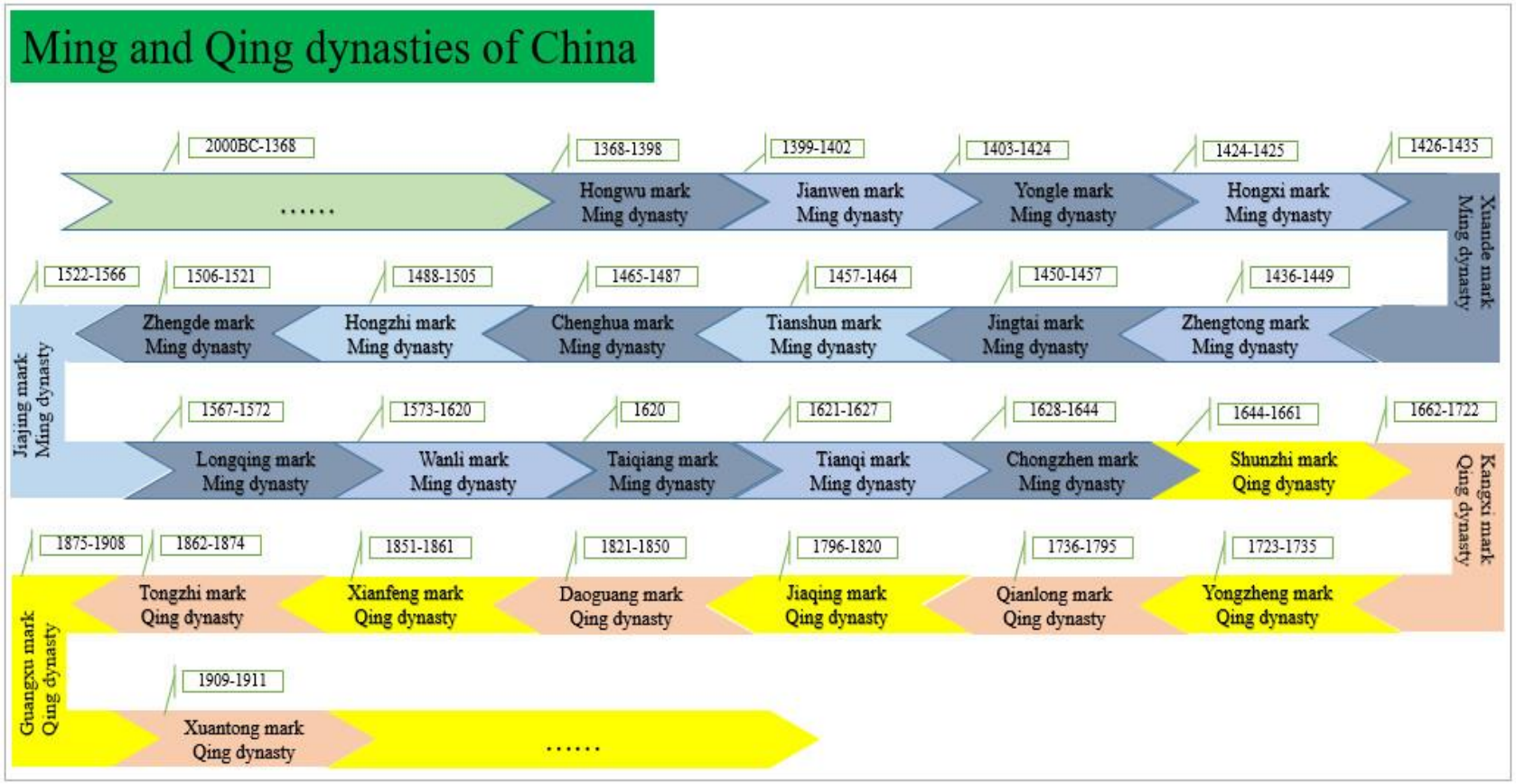
How to build multilingual terminological resources based on the domain ontology for experts and students' communication in the domain of Chinese ceramic vessels?

How to implement ontoterminology in Protégé, and in particular how to express essential characteristics in Protégé?

02

Domain and Objectives

Domain



When?

From 1368 to 1911

Objects?

Chinese ceramic vases

蒜头瓶
Suantouping; 'Garlic-Mouth' Vase;
'Garlic-Head-Shaped' Vase

Domain

Data collection of research

Criteria?

The objects selected:

- are as much different as possible from one another
- come from well-known collections of ceramic vessels in China
- the information on them is publicly available

故宮博物院
THE PALACE MUSEUM

Collections?

Museum	Number of objects
Palace Museum, Beijing	97
National Museum of China, Beijing	22
Guangdong Museum, Guangzhou	24
Shanghai Museum, Shanghai	4
Capital Museum, Beijing	2

Objectives

- Build a bilingual (Chinese and English) terminological knowledge base (e-dictionary) of Chinese ceramic vases for archeologists and students.
- Build an ontology to represent knowledge in the Chinese ceramic vases of Ming and Qing dynasties and publish these open linked data on the LOD.
- Propose an approach for translating essential characteristics into Protégé.
- Provide a reference for archaeologists, knowledge engineers, ontology engineers, and terminologists working on this domain
- Enrich existing methodologies of building domain ontology by means of a term-and-characteristic guided approach so as to reduce the dependency on logic and formal language

03

Theories and Research map

Theories

Term

Verbal designation of a general concept in a specific subject field (ISO 1087-1).

Characteristic

Abstraction of a property (ISO 1087-1).

Concept

Unit of knowledge created by a unique combination of characteristic (ISO 1087-1).

Essential characteristic

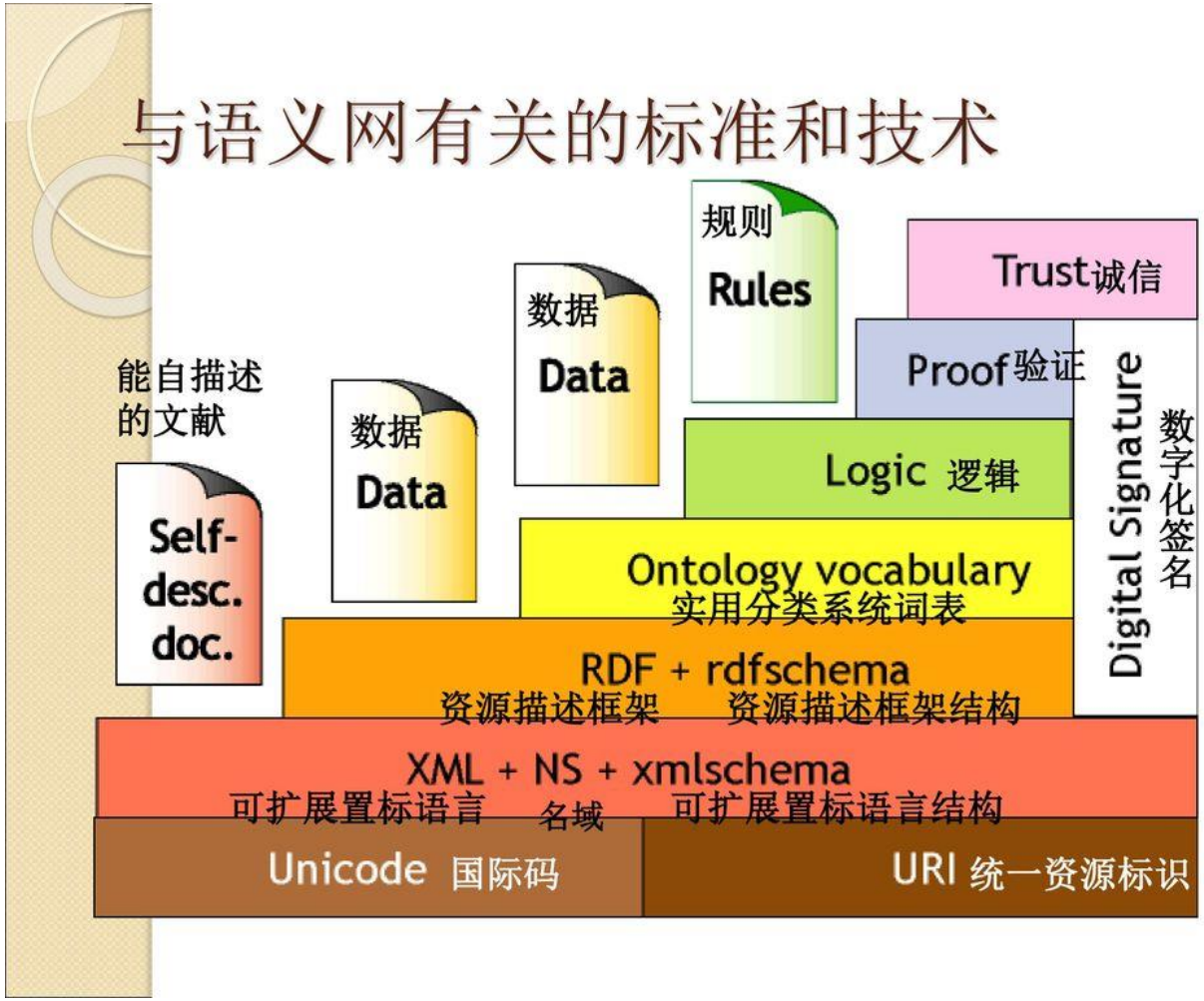
A characteristic indispensable in order to understand the concept (ISO 1087-1).

Theories

Semantic Web:

The Semantic Web provides a common framework that allows data to be shared and reused across application, enterprise, and community boundaries (W3C).

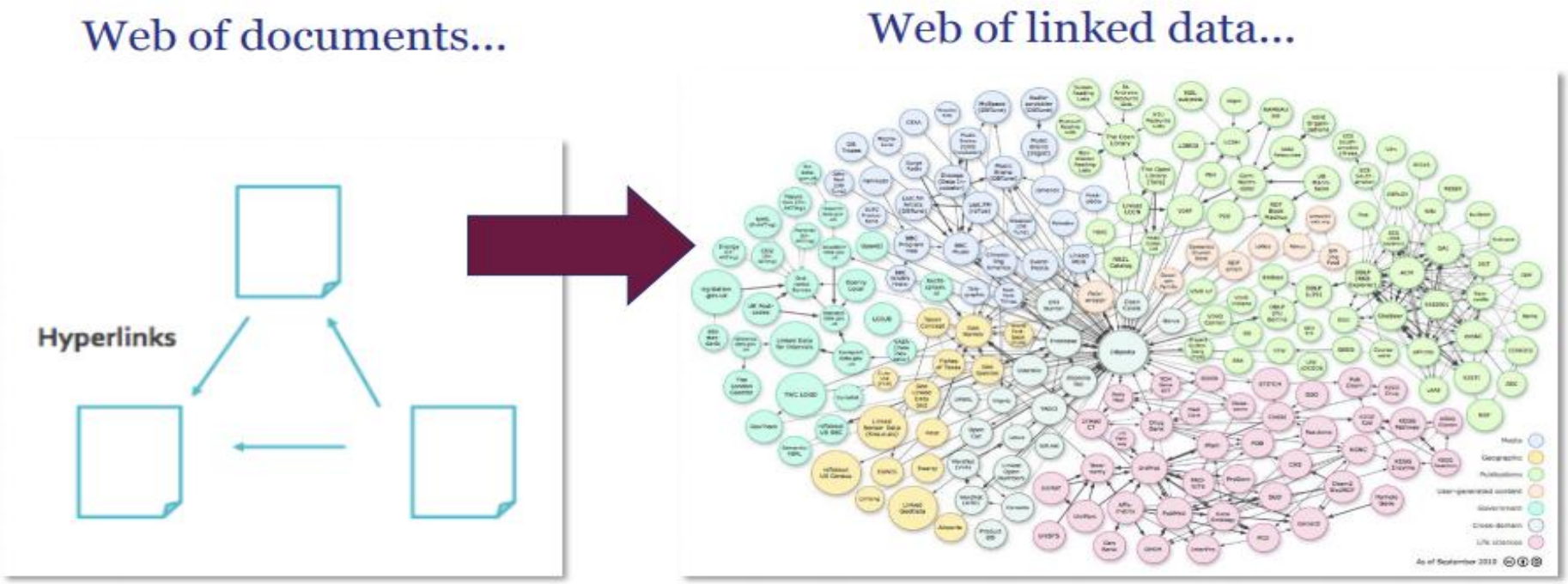
Ontology vocabulary and RDF are the core of the semantic web.



Theories

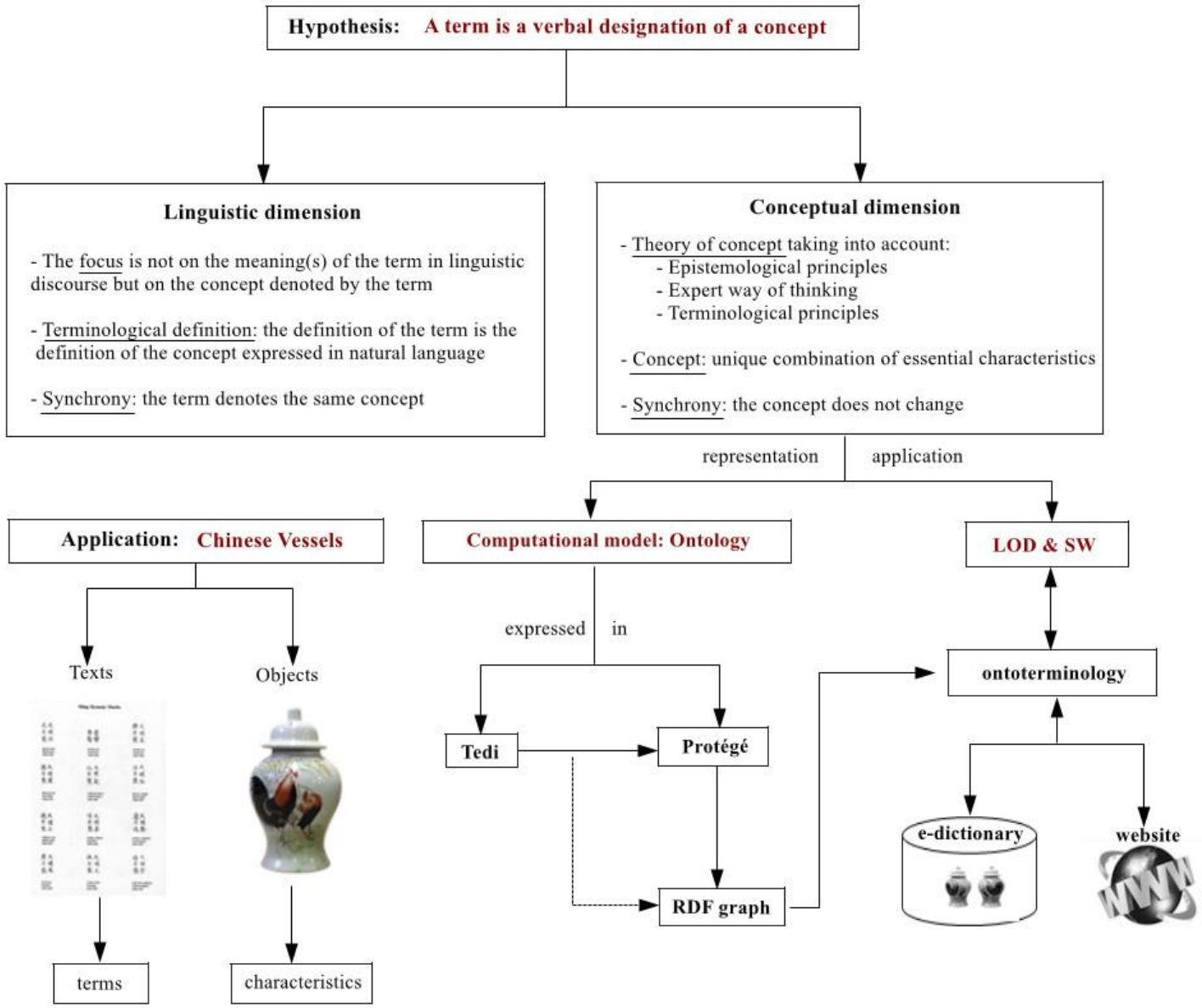
Linked data:

a term used to describe a recommended best practice for exposing, sharing, and connecting pieces of data, information, and knowledge on the Semantic Web using URIs and RDF (Tim Berners Lee, 2006).



Research map

Terminology work includes two dimensions – Linguistic dimension and Conceptual dimension.



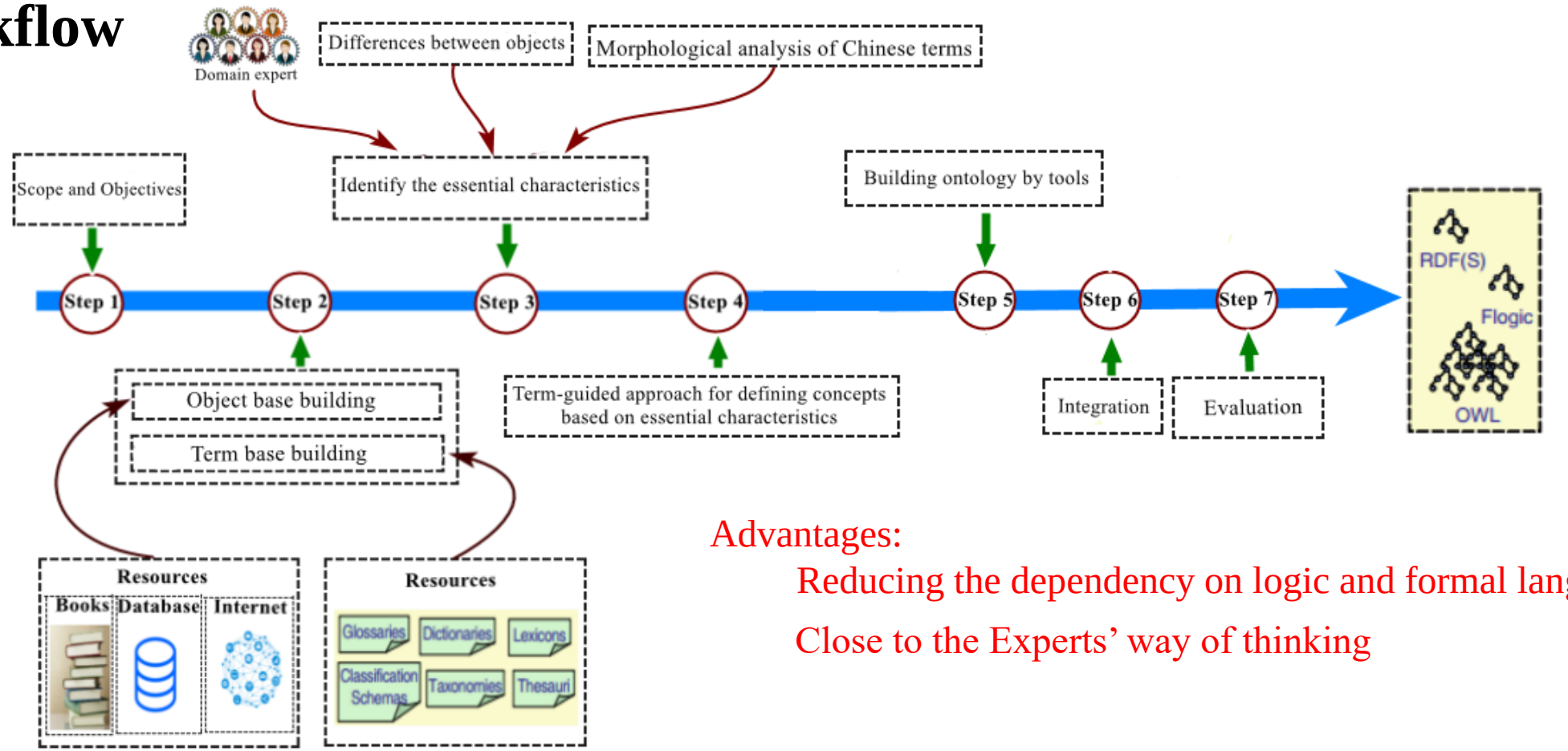
04

Methodology

Methodology

The term-and-characteristics guided methodology is partly based on ISO 1087 (Terminology work)

workflow



Advantages:

Reducing the dependency on logic and formal language

Close to the Experts' way of thinking

Methodology

Step 1. **Identify the scope of the domain and the objectives.**

This step is the beginning of building ontology. The aim is to define the scope of the project and its objectives. The competency questions are used for this purpose.

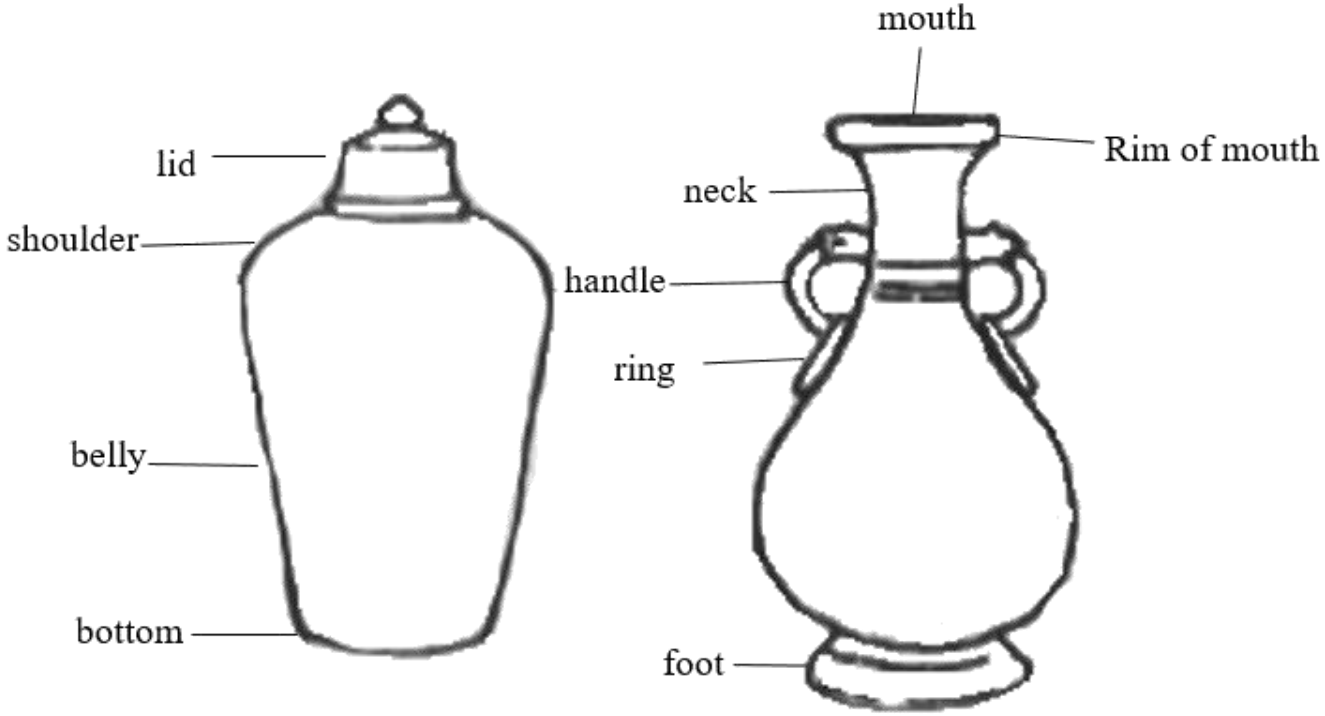
Step 2. **Identify terms and Objects**.

This stage aims at two goals. The first one is to select the set of vases to study from different museums. That reference set has to represent the richness of the domain without being too big (the primary goal of this project is to define the ontology and not to populate it).
The second one is to collect the set of terms corresponding to the selected vases.
Museum collections, web sites, databases, and books were the sources of information.

Methodology

Step 3: Identify essential characteristics.

we need to identify the essential characteristics on which the definition of concepts relies. The method of identifying essential characteristics includes **an object point of view** and **a term point of view**.

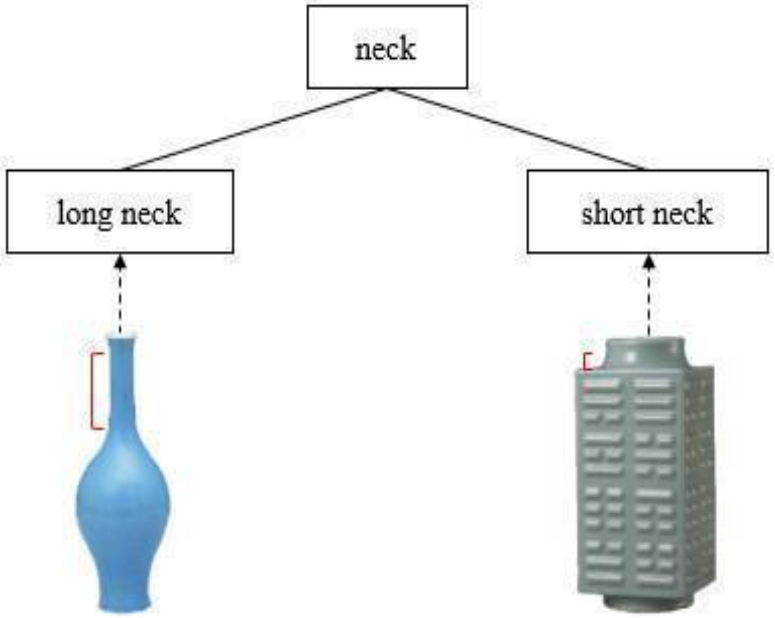


An object point of view.
Vase parts are named after human bodies

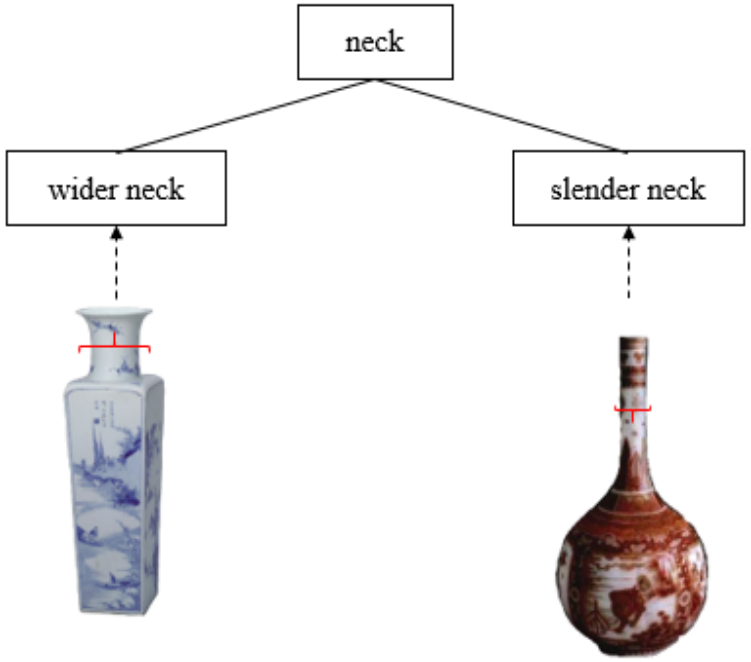
Methodology

Example： Neck part

Axis of analysis: Neck length



Neck width



Bending degree of the neck

Essential characteristics:
- /long neck/
- /short neck/

- /wider neck/
- /slender neck/

- /straight neck/
- /bending neck/

Methodology

A term point of view: **Morphological Analysis of Chinese Terms**

The Nanjing museum adopts the following order of modifiers for naming ceramic:
“dynasty + kiln + glaze + colour + decoration + shape + texture + type”.

The information conveyed by the modifiers expresses knowledge of different nature, either essential, such as shape, material, and type, or descriptive, like glaze and color.

For example, the term “清 雍正 粉青釉 凸花 如意耳 蒜头 瓷 瓶” conveys the descriptive characteristics of dynasty (“清” Qing dynasty), emperor (“雍正” Yongzheng mark), glaze-color (“粉青釉” powder blue glaze), and decoration (“凸花” designed with flowers).

It also conveys the essential characteristics of handle (“如意耳” Ru-Yi handle), shape (“蒜头” garlic-like head), material (“瓷” porcelain), and type (“瓶” vase).

Methodology

Step 4： A Term-guided approach for defining concepts based on essential characteristics. This approach relies on the fact that a concept is a set of essential characteristics that is stable enough to be named in a given natural language

Example:

The term “garlic-head vase” denotes the following set of essential characteristics:
{/vase/, /one mouth/, /garlic shape mouth/, /ring foot/}.

The name of the concept denoted by “garlic-head vase” is built from the set of essential characteristics:
<**Vase** one mouth garlic shape mouth ring foot>,

The definition of the term “garlic-head vase” is the translation in natural language of the formal definition of the concept: “Vase with a garlic shape mouth and ring foot”.

Methodology

The term "**garlic-head vase I**" denotes the following set of essential characteristics:
{/Garlic-headVase/, /short neck/, /circle shoulder/, /ru-yi shaped handle/, /globular belly/}.

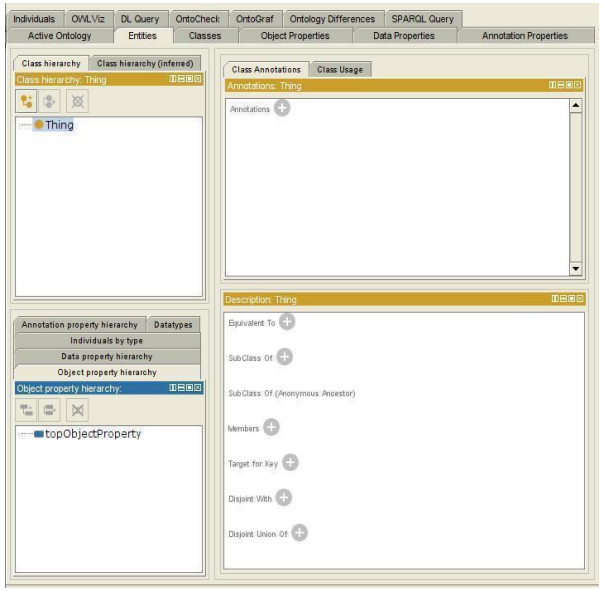
The concept name (is not a term) denoted by "garlic-head vase I" is:
<**Garlic-headVase** short neck circle shoulder ru-yi shaped handle globular belly>,

The definition of the term "garlic-head vase I" is "Garlic-head Vase with short neck, circle shoulder, ru-yi shaped handle, and globular belly".

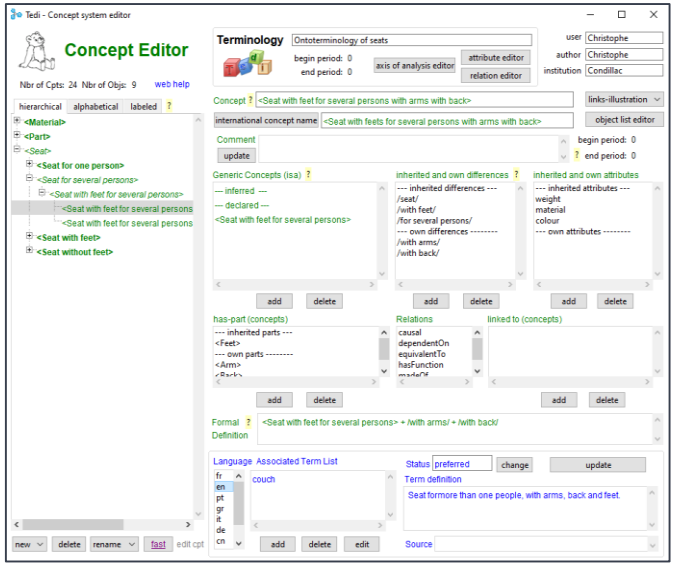
Methodology

Step 5：Building ontology by tools.

Protégé and Tedi differ on many points that could be summarized by saying that the former is as universal as the latter is specific.



Protégé vs Tedi



Protégé is a free, open-source software, the most widely used ontology editor, supported by a large community of users.
But it can not directly represent the essential characteristics in DL.
Linguistic dimension is reduced (in general) to annotations.

Tedi is intended for experts to build ontoterminologies in accordance with the ISO principles on Terminology.
It relies on essential characteristics.
Linguistic dimension explicitly represented.

Methodology

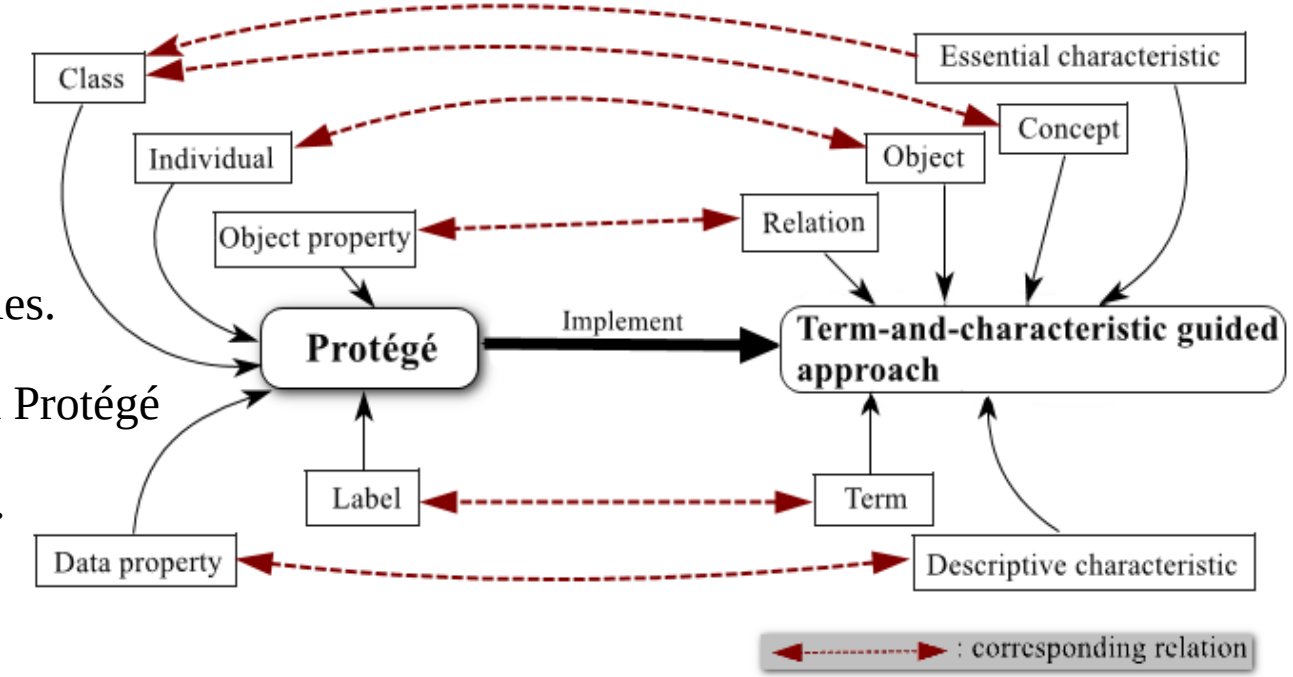
Step 5： Building ontology by tools.

Term: Terms are expressed as labels
(using annotated links such as rdfs:label)

Relation: Relations are represented as object properties.

Concept: Concepts are translated as a named class in Protégé

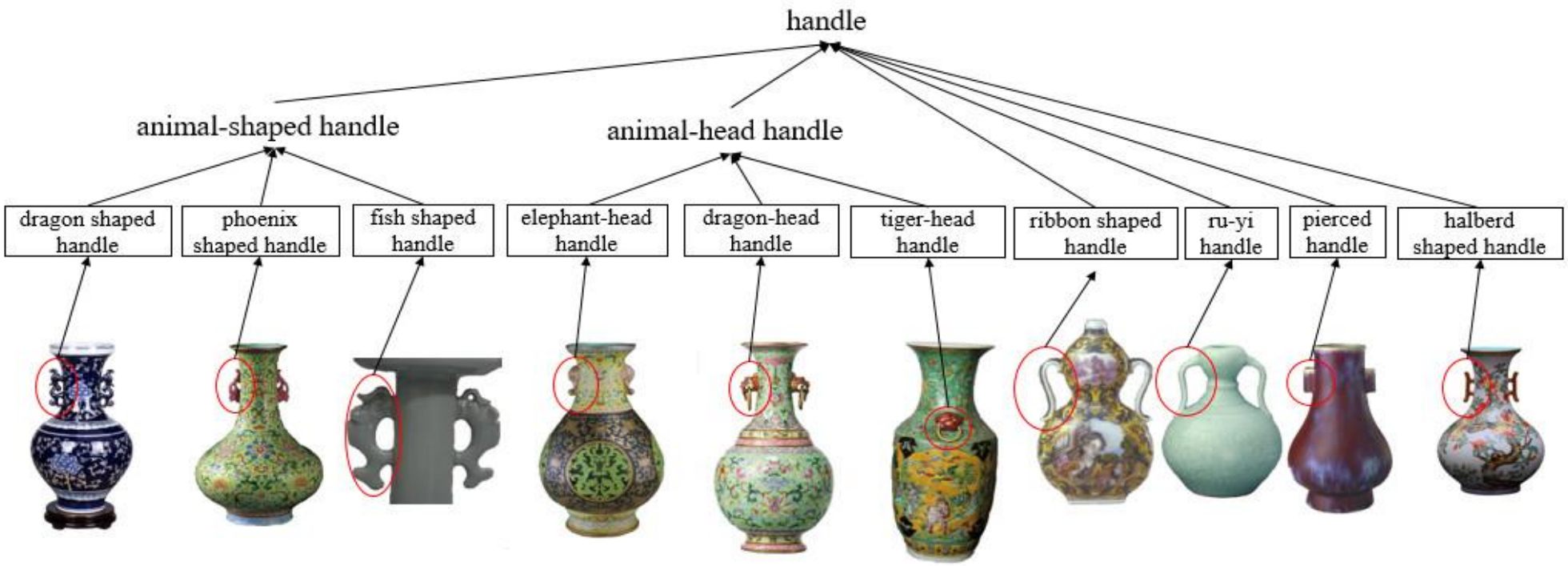
Object: Objects correspond to individuals in Protégé.



Descriptive characteristic: they are translated either as data properties if their value is a data literal or as object properties and classes if the value is an individual.

Essential characteristic: Since essential characteristics correspond to rigid predicates they cannot be directly expressed into Description Logic. Essential characteristics are expressed as classes.

Methodology



We defined essential characteristics as classes in order to be further sub-partitioned if necessary.

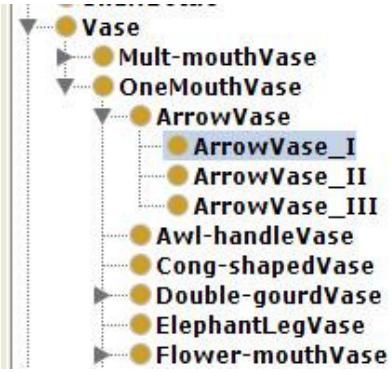
For example, the animal-head handle could be further sub-partitioned as elephant-head handle, dragon-head handle, and tiger-head handle.

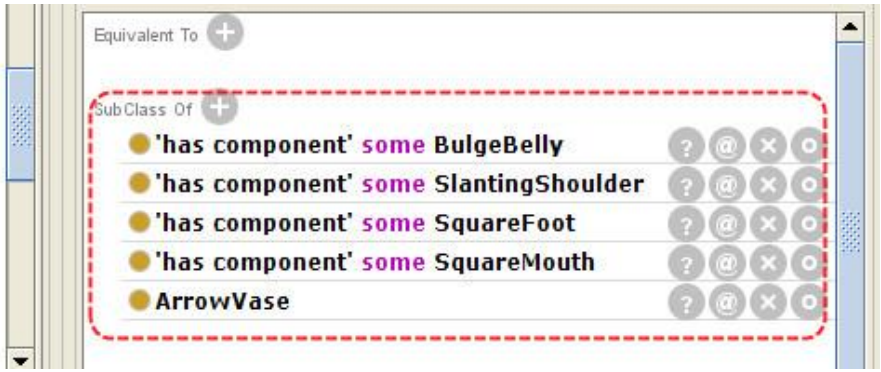
Methodology

Step 5：Building ontology by means of tools.

Owing an essential characteristic for a concept (class) is represented as a restriction of an object property whose range is the class associated to the essential characteristic.
This means that the concept (class) is a subclass of the anonymous classes defined by the property restrictions.

The concept denoted by the term “arrow vase I” includes the essential characteristics {/Arrow Vase/, /square mouth/, /slanting shoulder/, /bulge belly/, /square foot/}. It will be represented as a subclass of:
------ ArrowVase
------ has component some SquareMouth
------ has component some SquareFoot
------ has component some SlantingShoulder
------ has component some BulgeBelly





Methodology

Step 6: Integration.

Mapping

The class Vessel is linked to the CIDOC CRM class "E22 Man-Made Object" which comprises physical objects purposely created by human activity.

Ontology prefixes:

Prefix	Value
	http://www.dh.ketrc.com/otcontainer/data/OTContainer.owl#
OTContainer	http://www.dh.ketrc.com/otcontainer/data/OTContainer.owl#
crm	http://www.cidoc-crm.org/cidoc-crm/
dc	http://purl.org/dc/elements/1.1/
dcterms	http://purl.org/dc/terms/
foaf	http://xmlns.com/foaf/0.1/
owl	http://www.w3.org/2002/07/owl#
rdf	http://www.w3.org/1999/02/22-rdf-syntax-ns#
rdfs	http://www.w3.org/2000/01/rdf-schema#
schema	http://schema.org/
skos	http://www.w3.org/2004/02/skos/core#
xsd	http://www.w3.org/2001/XMLSchema#

Art & Architecture Thesaurus

Methodology

Step 7: Evaluation.

We used two online platforms and queried the ontology against the competency questions.

OOPS!: An online tool to detect some of the most common pitfalls appearing when developing ontologies.
OOPS! has detected only minor pitfalls for the TAO CI ontology.

Results for P08: Missing annotations.	165 cases | Minor
Results for P13: Inverse relationships not explicitly declared.	7 cases | Minor

OntoMetrics: An online platform to calculate more advanced ontology metrics.

Metric	Value
Attribute richness	0.048485
Inheritance richness	2.715152
Relationship richness	0.334324
Class/Relation ratio	0.245171
Average population	0.8
Class richness	0.321212

"a good ontology does not perform equally well with regard to all criteria"

(Denny, 2009)

Methodology

Competency questions

CQ	Competency Question	Class(es)	*Relation*
1	What are the different types of vases?	Vase	?vase *is-a* Vase
2	What material the vase is made of?	Vase, Material	aVase *isMadeOf* ?material
3	What is the glaze color of the vase?	Vase, GlazeColor	aVase *hasGlazeColor* ?glazecolor
4	Which dynasty is the vase?	Vase, Dynasty	aVase *hasDynasty* ?dynasty
5	Which emperor is the vase?	Vase, Emperor	aVase *hasEmperor* ?emperor
6	What are the Chinese and English terms of vases?	Vase	?vase *label* ?string
7	Which temperature was the vase fired at?	Vase, Temperature	aVase *isFiredAt* ?temperature
8	What are the components of a vase?	Vase, Component	aVase *hasComponent* ?component
9	What is the function of a vase?	Vase, Function	aVase *hasFunction* ?function
10	Which Dynasty does an Emperor belong to?	Emperor, Dynasty	aEmperor *belongTo* ?dynasty
11	What is the foot diameter of a vase?	Vase	aVase *diameterOfFoot* ?string
12	What is the height of a vase?	Vase	aVase *height* ?string
13	Which collection does a vase belong to?	Vase	aVase *collectedIn* ?string
14	Which kiln produced a vase?	Vase	aVase *producedIn* ?string
15	What is the decoration of a vase?	Vase	aVase *decoratedBy* ?string
16	What are the images of a vase?	Vase	aVase *image* ?string
17	What is the definition of a vase?	Vase	aVase *definition* ?string

Methodology

Q6: What are the Chinese and English terms of vases?

```
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
PREFIX otc: <http://www.dh.ketrc.com/otcontainer/data/OTContainer.owl#>
SELECT ?english_name ?chinese_name
            WHERE {  ?vase rdfs:subClassOf* otc:Vase.
                     ?vase skos:prefLabel ?english_name.
                     ?vase skos:prefLabel ?chinese_name.
                     FILTER (lang (?english_name)='en')
                     FILTER (lang (?chinese_name)='zh') }
ORDER BY ?english_name
```

english_name	chinese_name
"arrow vase"@en	"贯耳瓶"@zh
"arrow vase I"@en	"贯耳瓶 I"@zh
"arrow vase II"@en	"贯耳瓶 II"@zh
"arrow vase III"@en	"贯耳瓶 III"@zh
"awl-handle vase"@en	"锤把瓶"@zh
"cong-shaped vase"@en	"琮式瓶"@zh
"double-gourd vase"@en	"葫芦瓶"@zh
"double-gourd vase I"@en	"葫芦瓶 I"@zh

05

TAO CI ontology

TAO CI ontology

TAO CI ontology is open access at: http://www.dh.ketrc.com/otcontainer/data/OTContainer.owl

Ontology metrics:

Metrics	
Axiom	3121
Logical axiom count	1492
Declaration axioms count	327
Class count	165
Object property count	11
Data property count	8
Individual count	132
Annotation Property count	15
Class axioms	
SubClassOf	469
EquivalentClasses	0
DisjointClasses	176

TAO CI ontology

165 classes
9 top level classes

Class hierarchy:

- owl:Thing
  - Component
  - Dynasty
  - Emperor
  - Function
  - Glaze-Color
  - Material
  - ShapeOfBody
  - Temperature
  - Vessel

11 object properties

Object property hierarchy:

- owl:topObjectProperty
  - belongTo
  - dependOn
  - hasComponent
  - hasDynasty
  - hasEmperor
  - hasFunction
  - hasGlaze-Color
  - isComponentOf
  - isDependedOn
  - isFiredAt
  - isMadeOf

Data property hierarchy:

- owl:topDataProperty
  - diameterOfFoot
  - diameterOfMouth
  - height
  - isCollectedIn
  - isDecoratedBy
  - isProducedIn
  - numberOfBelly
  - numberOfMouth

8 data properties

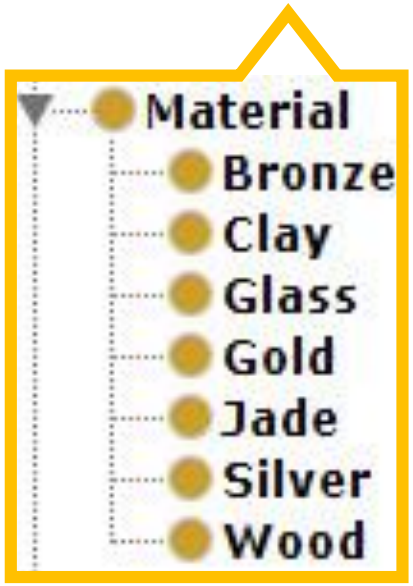
Individuals:

- arrow_vase_001
- arrow_vase_002
- arrow_vase_003
- arrow_vase_004
- arrow_vase_005
- arrow_vase_006
- arrow_vase_007
- arrow_vase_008
- ash_glaze
- awl-handle-shaped_vase_001
- blue-and-white
- blue_glaze
- bronze
- celadon_glaze
- Chenghua
- Chongzhen
- circle_rouleau_vase_001
- circle_rouleau_vase_002
- circle_rouleau_vase_003
- clay
- clear_glaze
- cloisonne
- cong-shaped_vase_001
- cong-shaped_vase_002
- cong-shaped_vase_003
- coral-red_glaze
- crystalline_glaze
- cylindrical_vase_001
- cylindrical_vase_002
- Daoguang
- dong_green_glaze
- double-gourd_shaped_vase_005
- double-tube_vase_001
- double_gourd_shaped_vase_001
- double_gourd_shaped_vase_002
- double_gourd_shaped_vase_003
- double_gourd_shaped_vase_004
- double_gourd_shaped_vase_006
- doucai
- dragon-massk_handle
- dragon_shaped_handle
- elephant_shaped_handle

132 individuals

TAO CI ontology

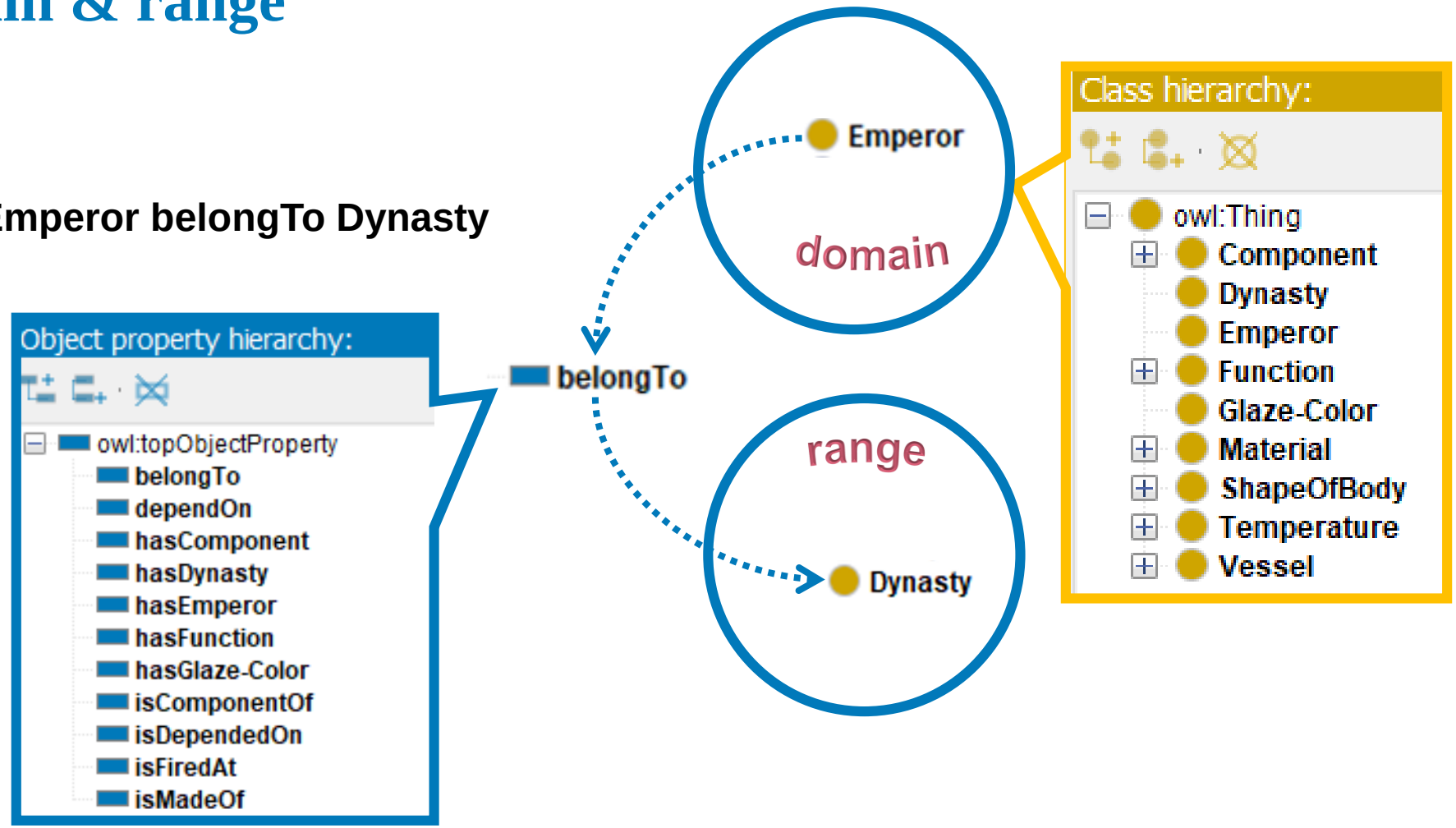
Material class



Different materials are the essential characteristics of Vessels, such as /bronze/, /clay/, /silver/, /glass/, /gold/, /jade/, /wood/. We defined them as classes in order to specify them if necessary. In order to be precise enough, we define them as disjointed.

Domain & range

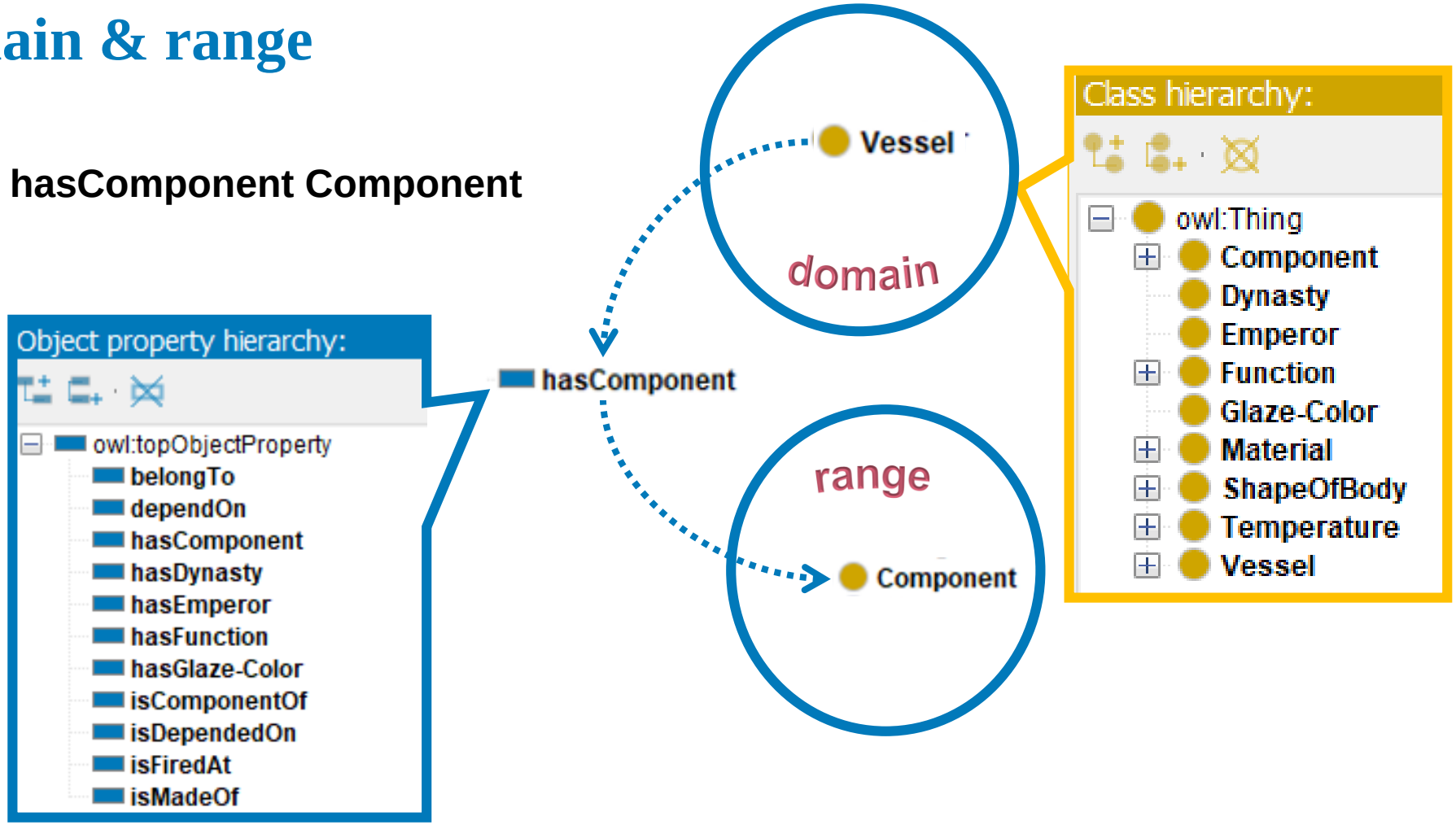
Emperor belongTo Dynasty



TAO CI ontology

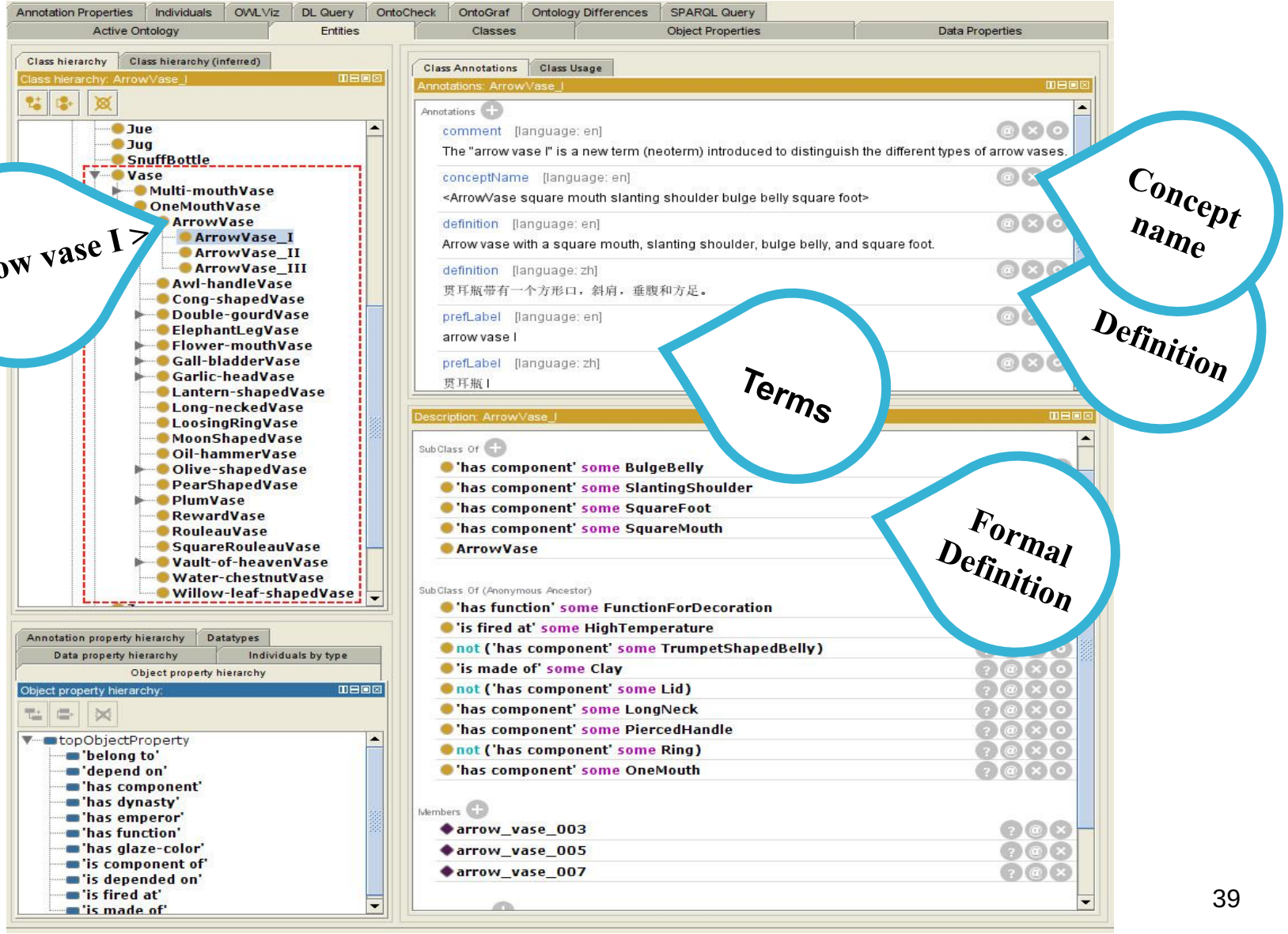
Domain & range

Vessel hasComponent Component



TAO CI ontology

The term “arrow vase I” designates the concept <Arrow Vase square Mouth slanting shoulder bulge belly Square foot>,
which is represented as the subclass of the following restrictions
------ ArrowVase
------ has component some SquareMouth
------ has component some SquareFoot
------ has component some SlantingShoulder
------ has component some BulgeBelly

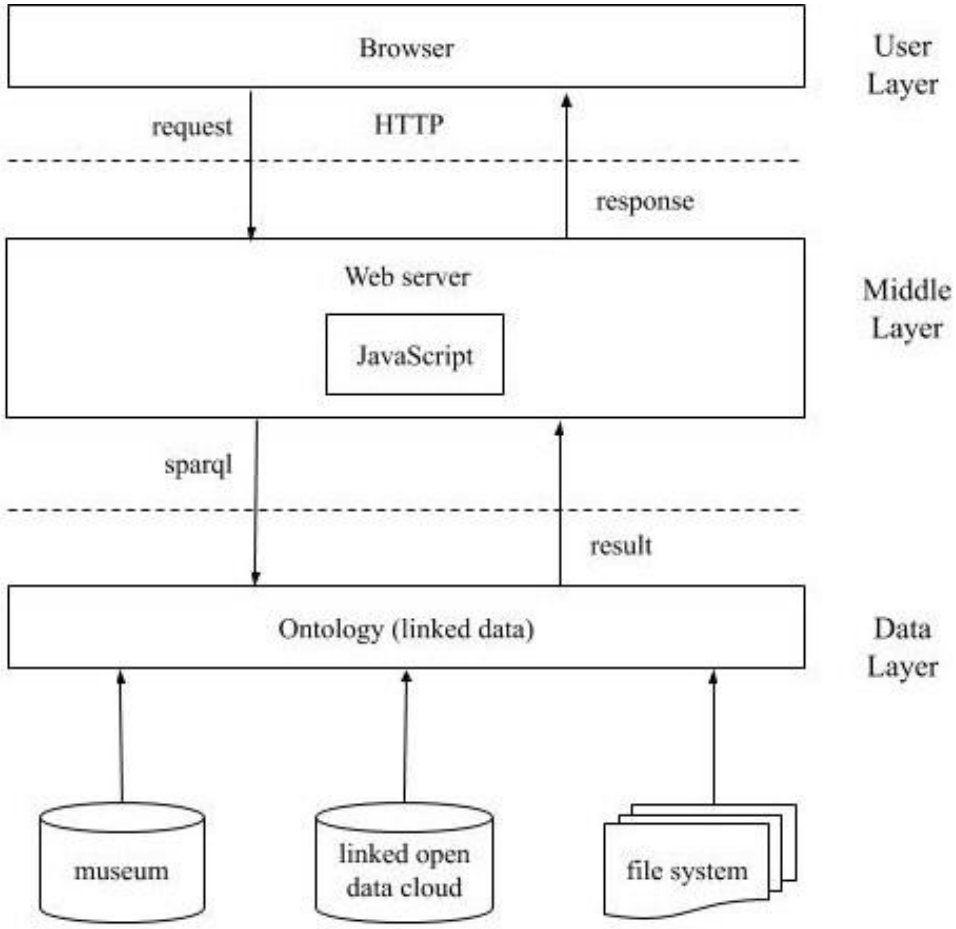


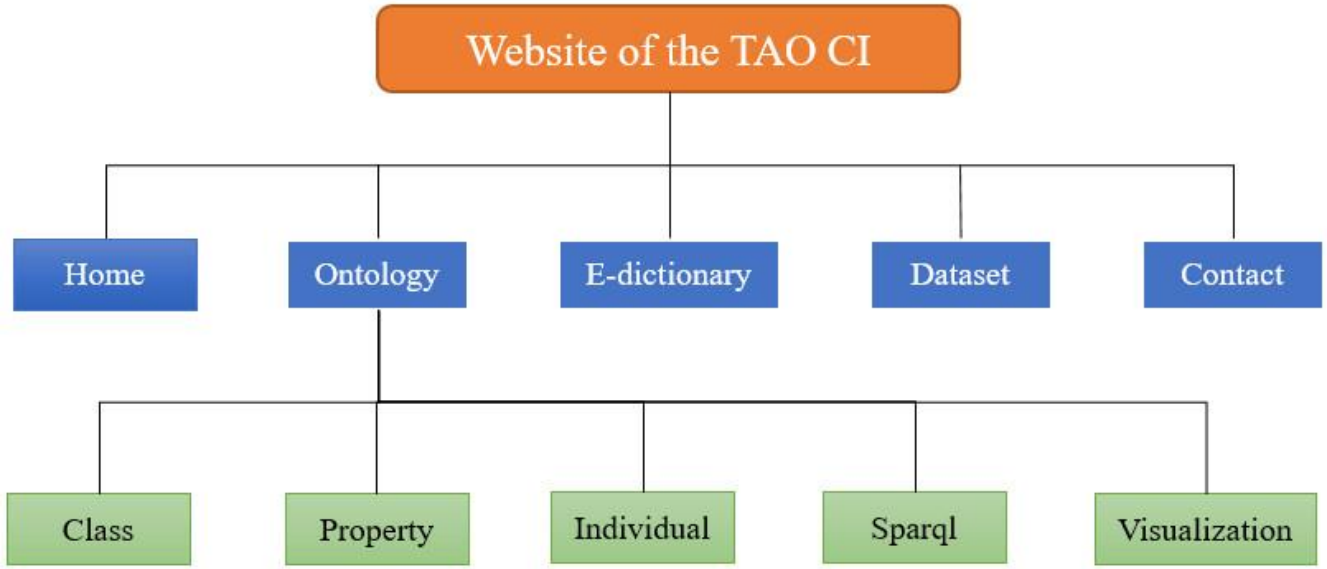
06

Website and e-Dictionary

Website

B/S structure, HTML+CSS+Javascript. Ontology provides the knowledge.





Function structure

Website

Please visit: http://www.dh.ketrc.com

e-Dictionary

JavaScript parses the ontology.

Arrow vase I

All information comes from the Tao Ci ontology.
Every term corresponds to a class in Tao Ci.

Arrow Vase
Arrow Vase I
Arrow Vase II
Arrow Vase III
Cong-shaped Vase
Double-gourd Vase
Double-gourd Vase I
Double-gourd Vase II
Double-gourd Vase III
Double-tube Vase
Elephant Leg Vase
Flower-mouth Vase
Flower-mouth Vase I
Flower-mouth Vase II
Gall-bladder Vase
Gall-bladder Vase I
Gall-bladder Vase II
Garlic-head Vase
Garlic-head Vase I
Garlic-head Vase II
Lantern-shaped Vase
Long-necked Vase
Loosing Ring Vase
Moon Shaped Vase
Oil-hammer Vase
Olive-shaped Vase
Olive-shaped Vase I
Olive-shaped Vase II
Pear Shaped Vase
Plum Vase

Arrow Vase I

Terms (prefLabel):
arrow vase I [EN]
贯耳瓶 I [ZH]

Concept Name:
<ArrowVase square mouth slanting shoulder bulge belly square foot>

Definition:
Arrow vase with a square mouth, slanting shoulder, bulge belly, and square foot.[EN]
贯耳瓶带有一个方形口，斜肩，垂腹和方足。[ZH]

Comment:
The "arrow vase I" is a new term (neoterm) introduced to distinguish the different types of arrow vases.

Essential Characteristic

- has component: /Square Foot/
- has component: /Bulge Belly/
- has component: /Slanting Shoulder/
- has component: /Square Mouth/
- is fired at: /High Temperature/
- has function: /Function For Decoration/
- is made of: /Clay/

See also
https://www.dpm.org.cn/collection/ceramic/227009.html
http://www.zhejiangmuseum.com/zjbwg/collection/collect_detail.html?id=3242

Terms in English & Chinese

Essential Characteristics

Individuals

07

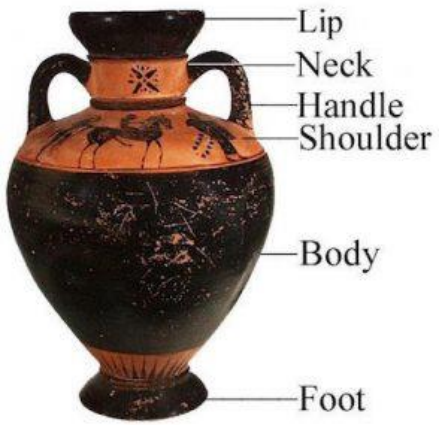
Conclusion and Future work

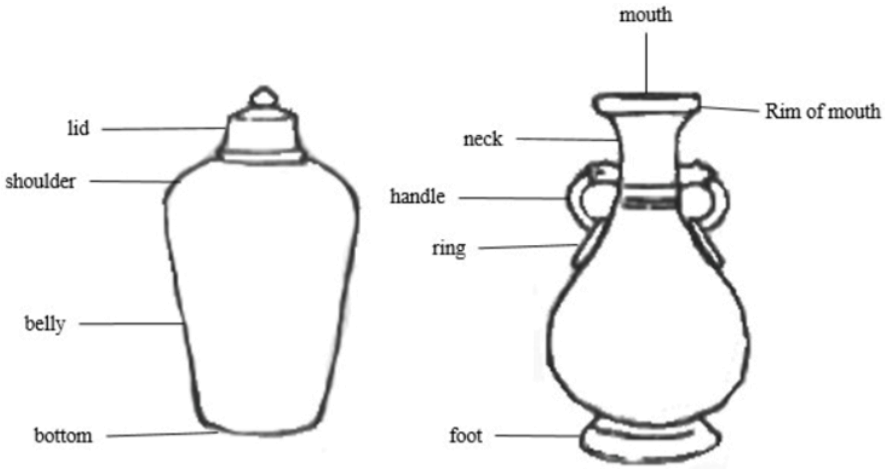
Conclusion

The objectives of this work have been successful achieved:

1. Bridge an existing gap by building an ontology to represent knowledge in the Chinese ceramic vases of Ming and Qing dynasties and publish these open linked data on the LOD cloud.
2. Build a bilingual (Chinese and English) terminological knowledge base (e-dictionary) of Chinese ceramic vases for archeologists and students.
3. Enrich existing methodologies of building domain ontology by means of a term-and-characteristic guided approach so as to reduce the dependence on logic and formal language.
4. Provide a reference for archaeologists, knowledge engineers, ontology engineers, and terminologists working on this domain.
5. Propose an approach for translating essential characteristics into Protégé.

Future work





Future work

Ontological dimension

- Complete the TAO CI ontology by considering other type vessels, such as Jar, Bowl, and Cup.

Linguistic dimension

- Import the OntoLex-Lemon model to enrich the linguistic dimension information.

Core ontology

- Proposed a core ontology of ceramic vessels based on the TAO CI ontology and Lekythos project.

Thanks.

Acknowledgements:

A special thanks for funding of the China and her expert advice to Effie Meramveliotaki, Archaeologist-Curator, Byzantine & East Asian Studies

The authors wish to gratefully acknowledge all the reviewers and my supervisors.

China Scholarship Council