Toward a Novel Application of CIDOC CRM to Underwater Archaeological Surveys

Olivier Curé[(1)], Mariette Sérayet[(2)], Odile Papini[(2)] , Pierre Drap[(2)]

Université Paris Est, France [(1)]
LSIS Marseille, France [(2)]

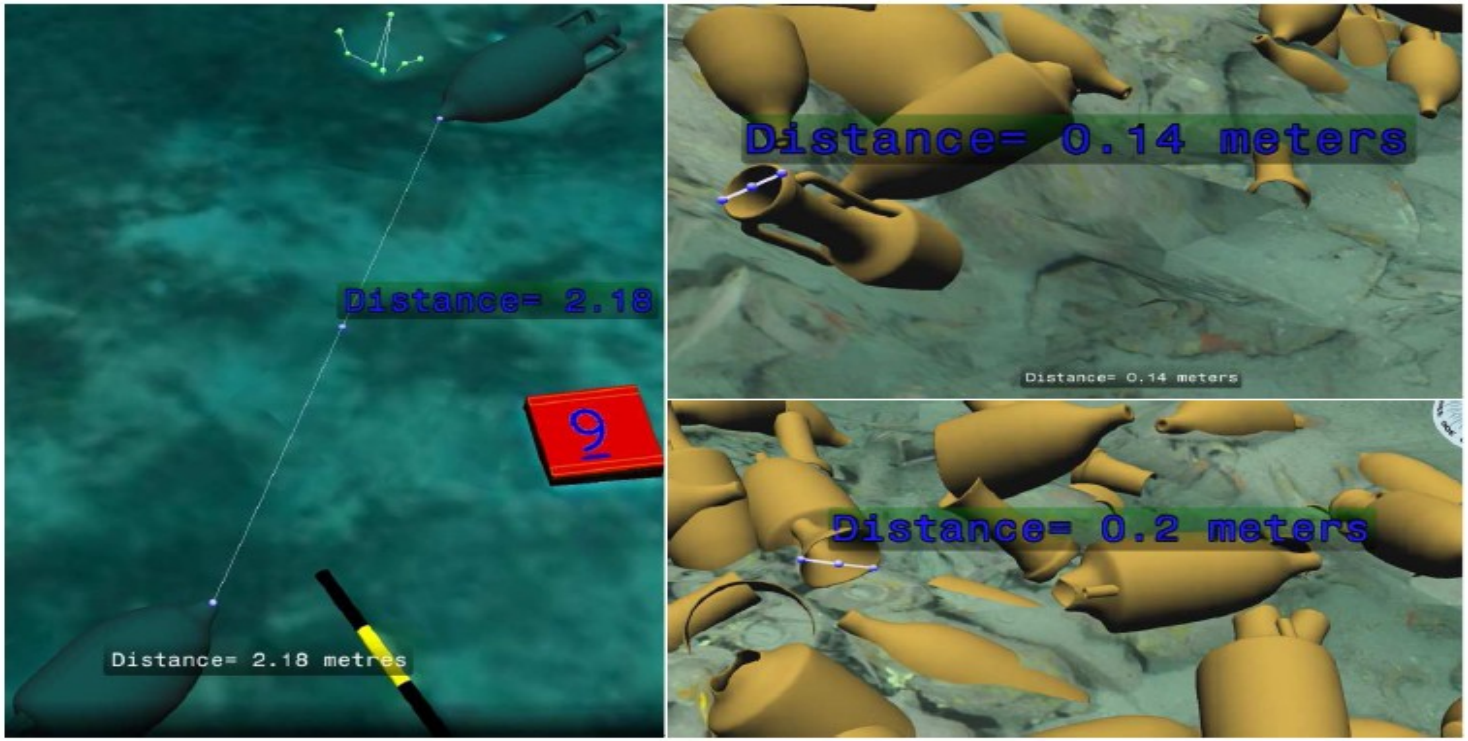
SWARCH-DL workshop @ ICSC 2010

Context

- The VENUS European project (FP6) aims to provide a virtual exploration of deep underwater archaeological sites.
- Virtual exploration of underwater environments will permit both experts and the general public to study interesting archaeological sites in a safe, cost-effective and pedagogical environment.

Objectives

- Represent and archive the digital artifacts corresponding to the studied items.
- A semantically rich representation requires an interpretation of archaeologists based on photogrammetry measures.



Objectives (2)

- Moreover, some non-observable features require inference rules to deduce missing values.
- Representation and inferences need to handle the uncertainty, inaccuracy and impreciseness of archaeological information.
- A knowledge base is needed.

Knowledge base

- Represented in Description Logics (DL):
  - For interoperability reasons
  - Adapted to represent the hierarchies of studied items (e.g. amphorae)
  - Sound and complete inferences performed by efficient reasoners, e.g. consistency checking.
  - Expressive
  - Interest of mapping to a cultural heritage ontology (i.e. ISO CIDOC CRM)

ICOM-CIDOC CRM

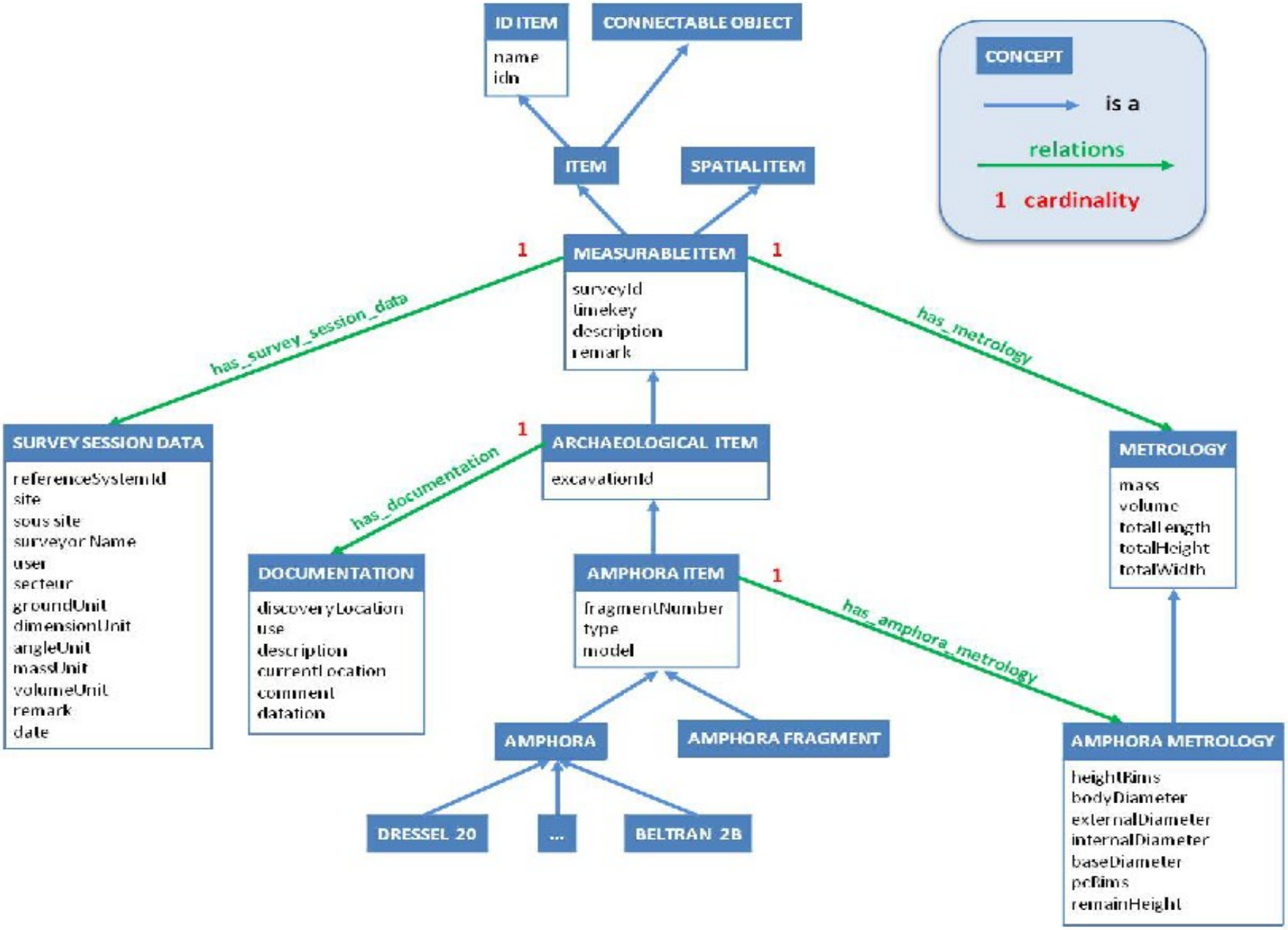
Committee on Documentation of the International Council of Museums Conceptual Reference Model

- CRM is an object oriented conceptual model which finds applications across the broader cultural heritage field.
  - Dedicated to the representation of any kind of archive or heritage piece of work.
  - Can be used to represent the way items are collected, identified and assigned with various more or less imprecise measurements, hypothetical attributes concerning their age and origin.

Ontologies

- Definition of an application ontology consisting of :
  - Archaeological knowledge is captured by a domain ontology (describing the vocabulary of studied items, e.g. amphorae).
  - Photogrammetry knowledge is represented within a task ontology
- Formalized in OWL2
  - Enables to represent for simple forms of constraints (e.g. on data values of certain items) and reason about concrete domains.

Application ontology



Mapping to CIDOC CRM

- A study of both ontologies emphasizes overlap and discrepancies:
  - Concepts and roles of our application ontology can be represented using CRM's TBox elements.
  - Concerning discrepancies, CRM is extended with elements form the application ontology TBox.
- Mapping is also expressed in DL

CIDOC CRM in OWL

- Definition of the E19_Physical_Object concept:
  - A sub class of E18_Physical_Thing that has a nuber of parts (P57) and has a current location (P55)
  - RDFS comment helps more:
  - *Scope Note: This class comprises items of a material nature that are units for documentation and have physical boundaries that separate them completely in an objective way from other objects.*

    The class also includes all aggregates of objects made for functional purposes of whatever kind, independent of physical coherence, such as a set of chessmen. Typically, instances of E19 Physical Object can be moved (if not too heavy). ...

Mapping

- Correspondences are defined using equivalence and subsumption relationships.
- This mapping has been defined manually, following a given methodology for the following reason:
  - Most of the semantics of the concepts and roles of CRM are defined in full text with rdf:comment properties.
  - Hence end-user needs to interpret these definitions.

Mapping example

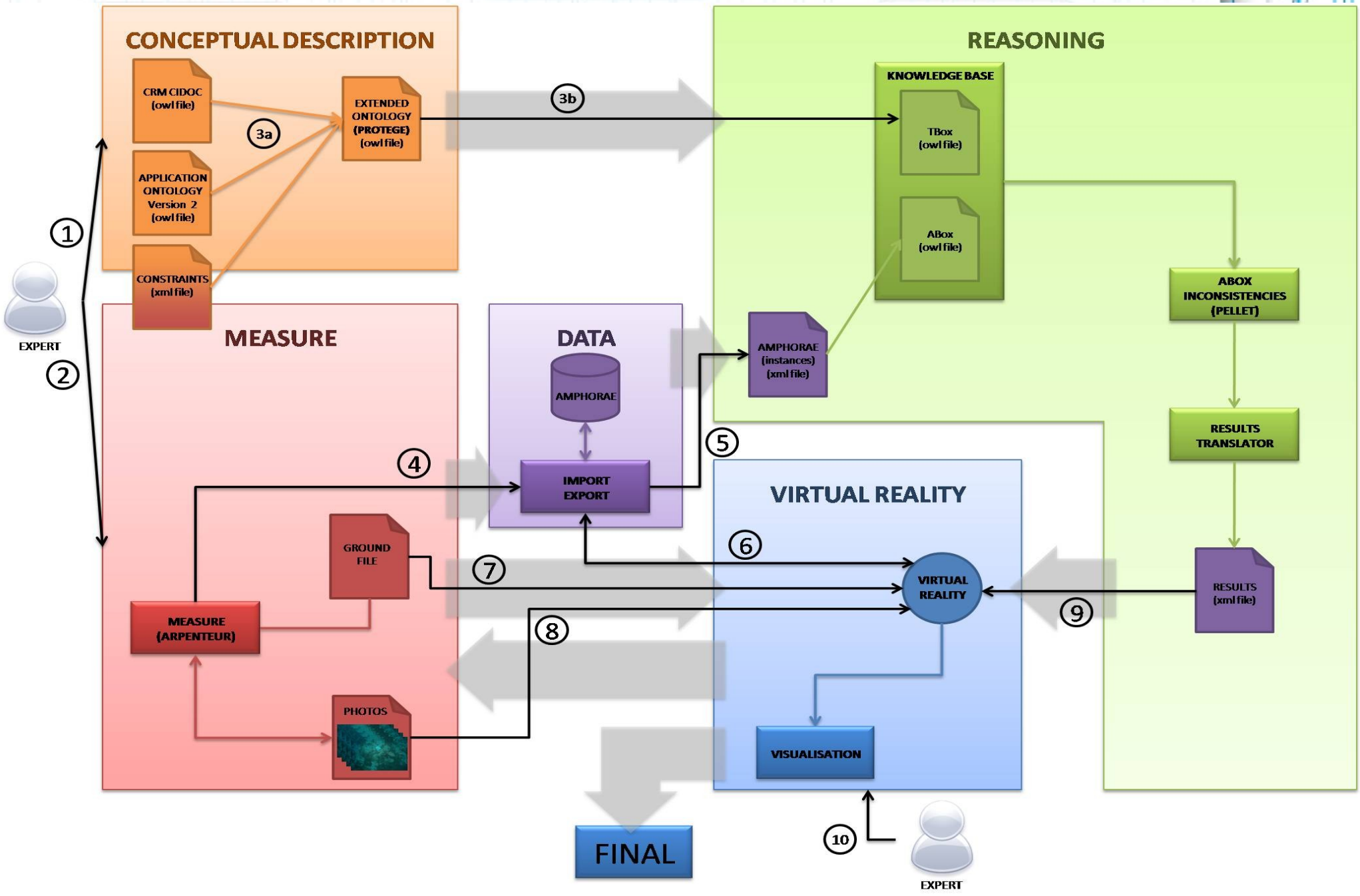
$$E\_22 \equiv MEASURABLE\_ITEM$$
$$E\_16 \sqsupseteq METROLOGY$$
$$E\_19 \equiv ID\_ITEM \sqcap CONNECTABLE\_ITEM \sqcap ITEM \sqcap SPATIAL\_ITEM$$
$$P\_48 \equiv idn$$
$$\top \sqsubseteq \forall hasSurveySessionData.E\_22$$
$$\top \sqsubseteq \forall hasSurveySessionData^{-}.E\_7$$

With:
E_22 : Man made object
E_16 : Measurement, the actions of measuring a physical objects)
E_19 : Physical object
P_48 : has preferred identifier property

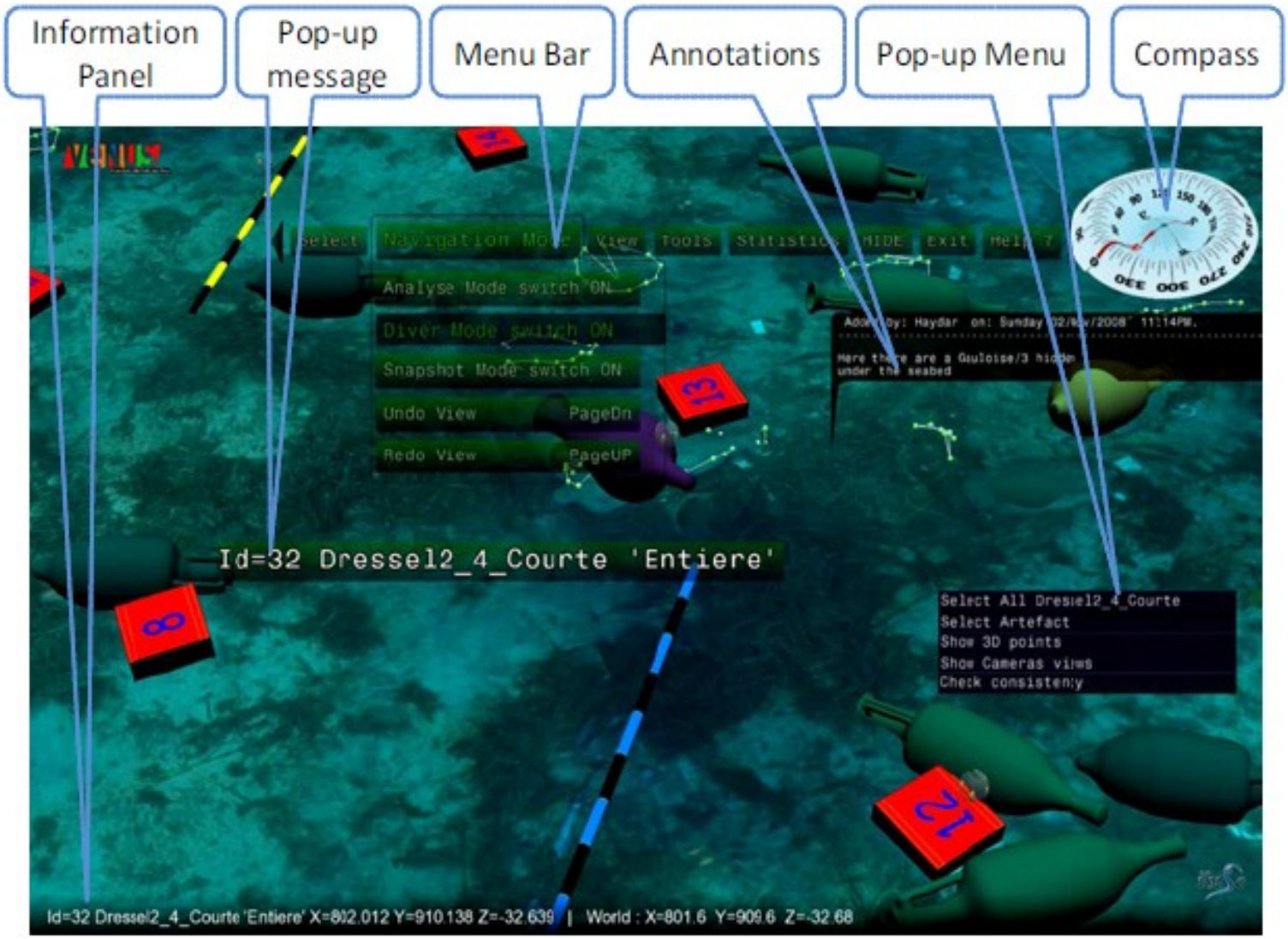
Reasoning with the merged ontology

- In our context, main of reasoning is to ensure consistency of the knowledge base:
  - At the ontology design stage: this is due to merging, extending the application ontology. This is performed using a DL reasoner within the Protégé editor.
  - At run-time: the ABox is responsible of inconsistencies. This requires to reason about the domains of properties. Repairing involves archaeologists through a graphical user interface.

Overview of reasoning



Visualization



With Augmented Reality

Conclusion and future works

- We proposed an extension of CIDOC CRM to archaeological knowledge
- Future works:
  - Handle uncertainty
  - Representation of expressive constraints such that enhancement of the overall data quality can be (semi) automatically performed.
  - Restoring consistency by accepting or refusing parts of the registered knowledge

Thank you for your attention

Questions ?