Setting up a CIDOC CRM Adoption and Use Strategy

CIDOC CRM: Success Stories, Challenges and New Perspective

George Bruseker

CIDOC 2017

Tblisi, Georgia

27/09/2017

Researcher,
Interpreter

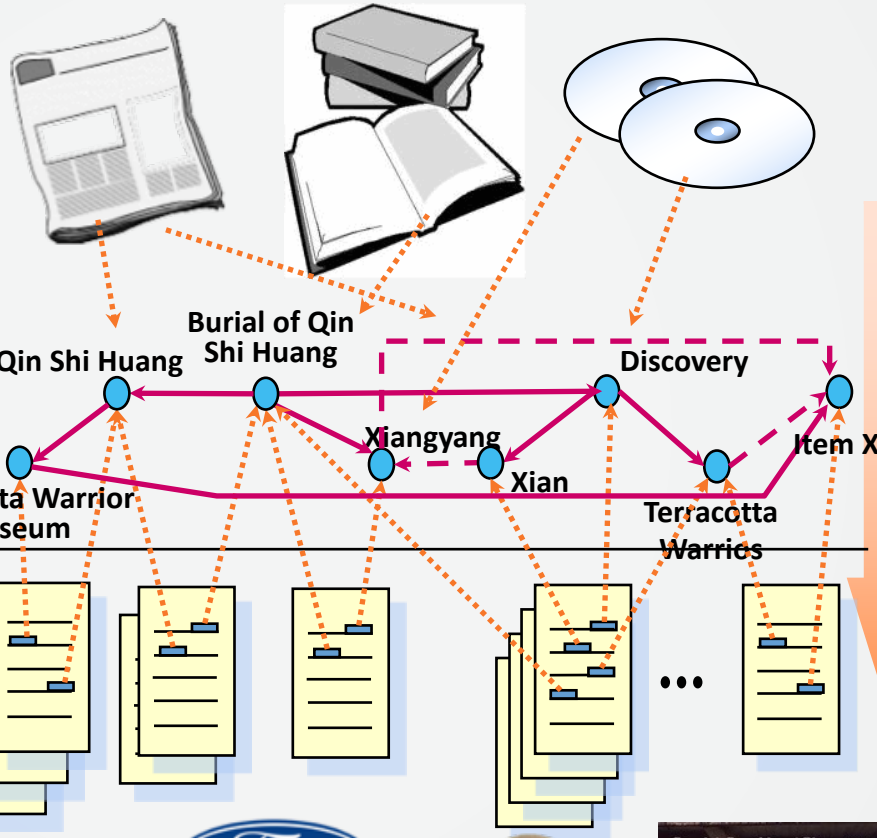
Goal: A Semantic Knowledge Ecosystem

Refer
interpret
present

Search,
correlate,
integrate

discover
collect
aggregate
update

reexamine

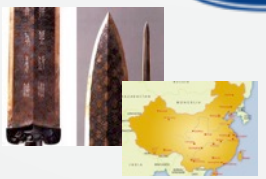


Publications
Stories
exhibitions

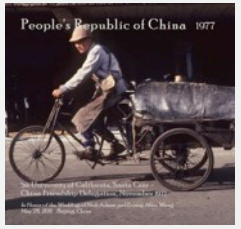
Layer of "Latest Knowledge"

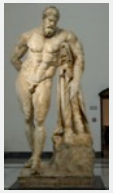
"Evidence layer"
Things
Sources
Collections
Corpora

Curator,
Conservator,
Excavator

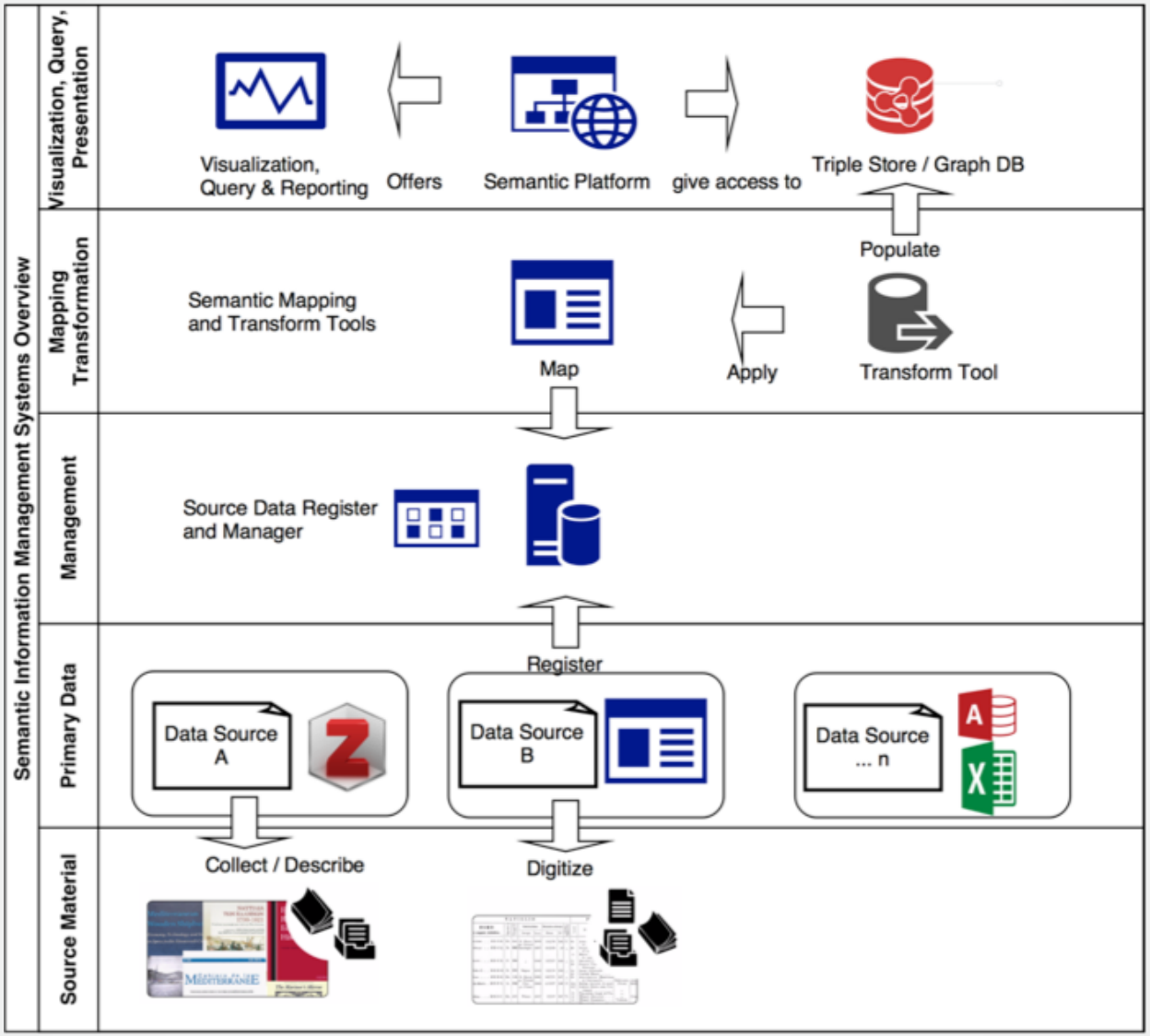


秦

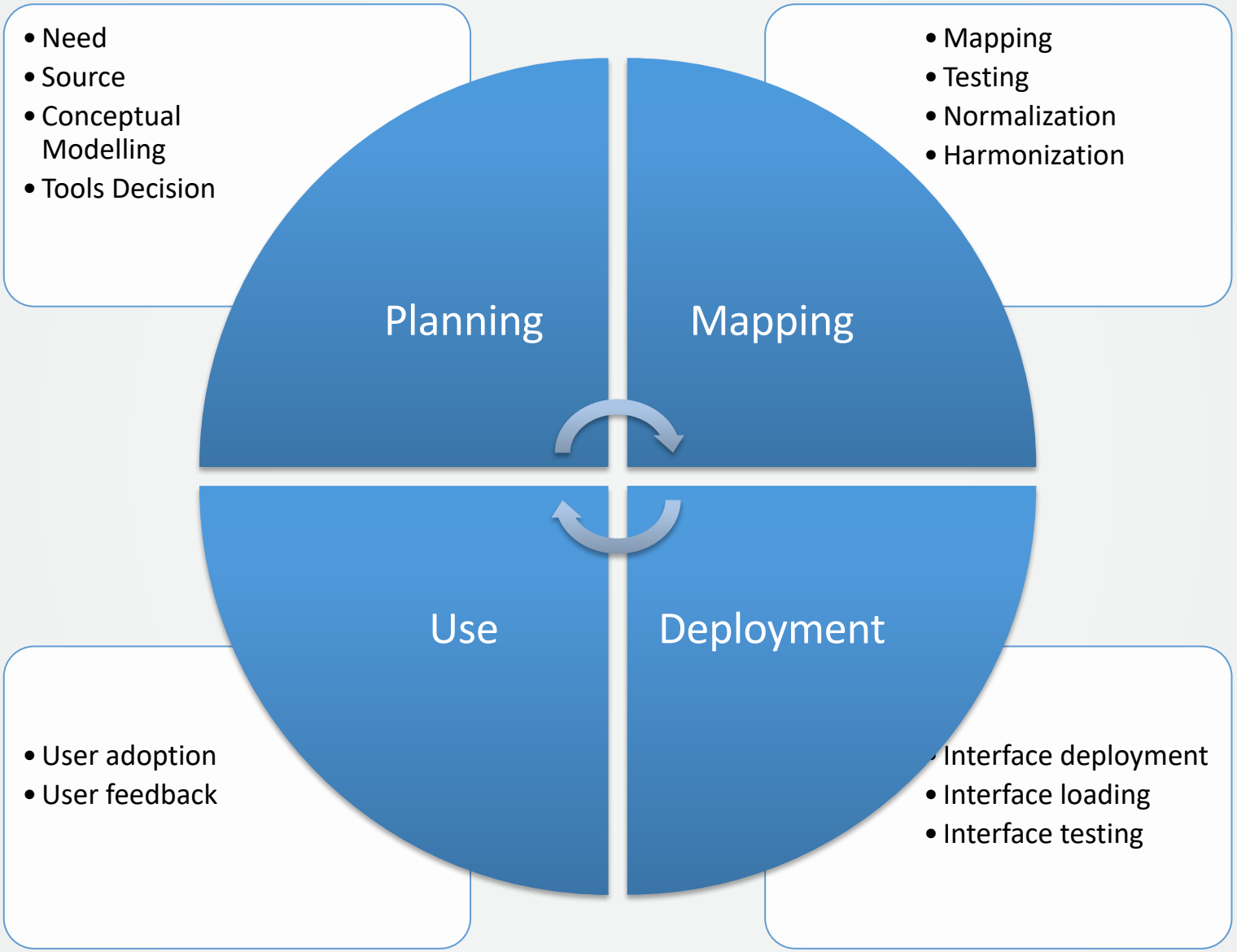




Components View of a CRM / Semantic Adoption Strategy



CRM Adoption Strategy Overview



Planning Overview

Planning	Mapping	Deploying	Use

1) Identify Need

2) Identify Sources

3) Conceptual Modelling

4) Identify Tools

Need Assessment

Planning	Mapping	Deploying	Use
Need	Sources	Modelling	Tools

IF

- Common research or presentation goals

AND

- Heterogeneous Data

AND

- One System not possible/desired

THEN

- **You need semantics!**

Source Analysis

Planning	Mapping	Deploying	Use
Need	Sources	Modelling	Tools

- List Sources
- Classify Sources (encoding/purpose)
- Identify Maintainers of Resources
- Identify who talks to who, who uses who's information?
- Study Semantic Content of Sources
- Identify and List Research Questions
- Choose most likely appropriate target ontologies
- Learn ontology

Tip: Data is in constant state of change. Use at least a spreadsheet to track source information over time.

Conceptual Modelling

Planning	Mapping	Deploying	Use
Need	Sources	Modelling	Tools

- Test target ontologies by attempting mappings
- Look for potential gaps in ontology constructs
- Use list of questions identified in source analysis to critique mapping
- Identify potential gaps in model and propose extension classes/relations suitable to your end

Tip: CRM is supported by a community of researchers and professionals. Ask the CRM list.

Tool Selection

Planning	Mapping	Deploying	Use
Need	Sources	Modelling	Tools

- You will need at least
  - Mapping tool
  - Triple Store or Graph DB
- You will probably want
  - Graphical Environment/platform for exploration and/or creation and/or visualization and reporting of semantic data
- Your decision will be affected by
  - Budget
  - Available expertise

Tip: One of the foundational points of CRM data is to unbind users from proprietary formats and commitment to closed data models. Consider always having an open transform of your data to RDF or OWL.

A Selection of Tools

Planning	Mapping	Deploying	Use
Need	Sources	Modelling	Tools

Sources Register

PARTHENOS

Pooling Activities, Resources and Tools for Heritage E-research Networking, Optimization and Synergies

Mapping Tools

Karma

Triple Stores / Graph DBs

OrientDB

neo4j

blazegraph

GraphDB 8

Semantic Data Management Platforms

WissKI

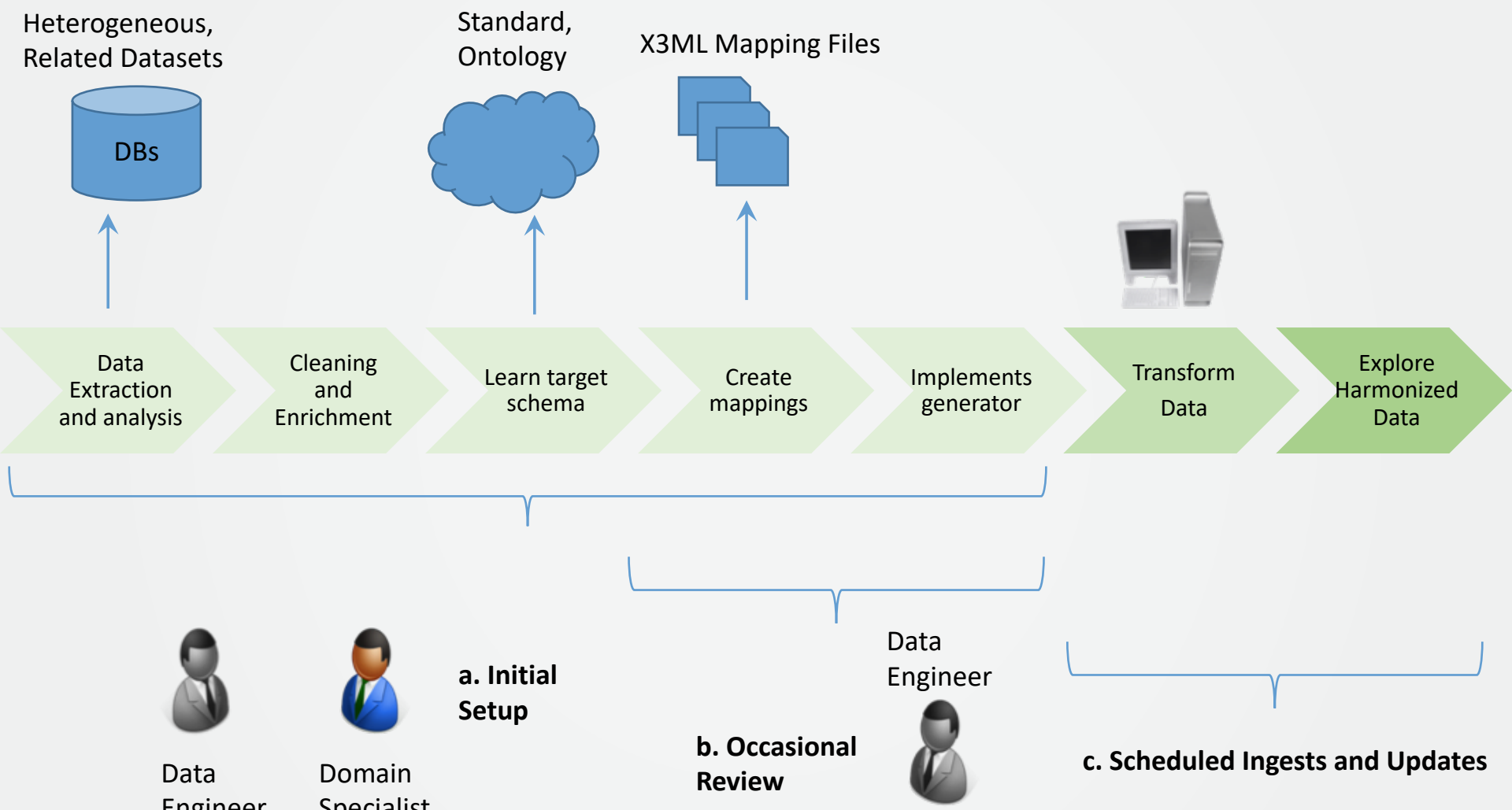
ResearchSpace

Mapping Overview

Planning	Mapping	Deploying	Use

- Mapping
- Testing
- Normalizing
- Harmonizing

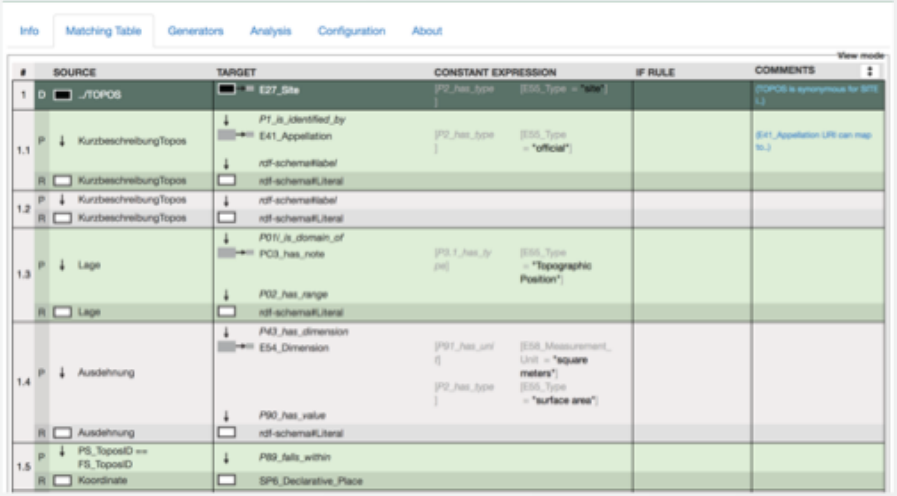
Workflow of Mapping Based on X3ML Suite



Mapping

Planning	Mapping	Deploying	Use
Mapping	Testing	Normalizing	Harmonizing

- Carried out by Domain Expert
- Carried out by someone who knows the source data
- Mapping takes place relative to research problematic documented in planning
- Mapping process should be saved, repeatable
- Mapping can be linked to dialogue with other partners in project

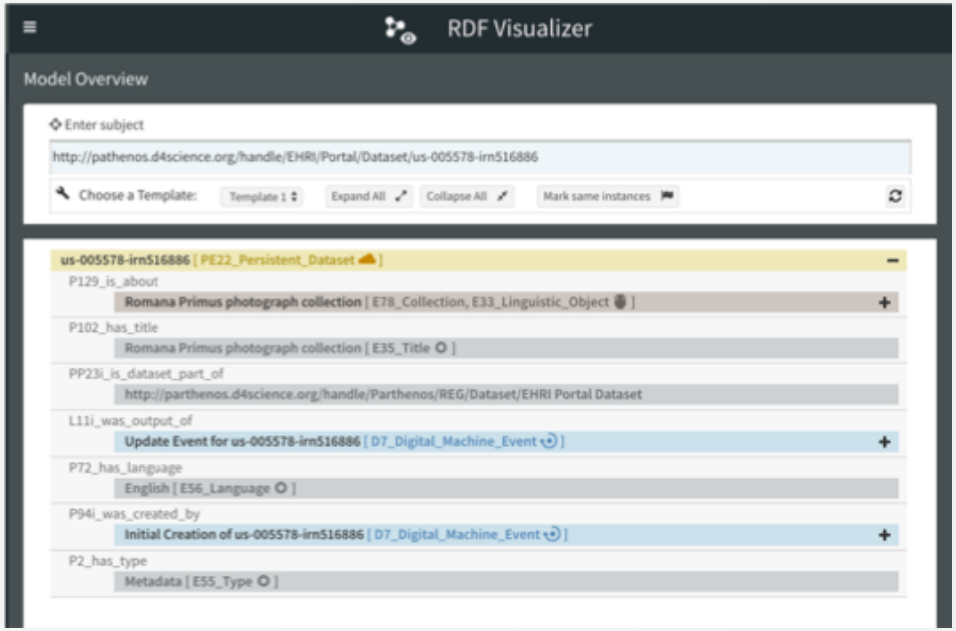


Screenshot X3ML Mapping Suite

Testing

Planning	Mapping	Deploying	Use
Mapping	Testing	Normalizing	Harmonizing

- Semantic mappings must ultimately ‘make sense’ to a human user
- Carried out by Domain Expert with Developer
- Check RDF or use visualization tool to verify transformed data communicates intended meaning
- Return to mapping where necessary
- Repeat



Screenshot RDF Visualizer

Data Normalizing

Planning	Mapping	Deploying	Use
Mapping	Testing	Normalizing	Harmonizing

- Schema to Schema matching from a source to CRM leaves data values unchanged.
- Normalization of data values according to international standard thesauri ensures better interoperability.
- Normalization Operation should be documented.

Data Harmonization

Planning	Mapping	Deploying	Use
Mapping	Testing	Normalizing	Harmonizing

- Ontology is like a pidgin language
- Like any language there are multiple ways to say the same thing
- After mapping individual sources, resulting maps must be semantically harmonized.
- Having a database of mappings allows you to manage this process.

Screenshot from X3ML toolsuite;
Db of mappings helps organize harmonization

Deployment Overview

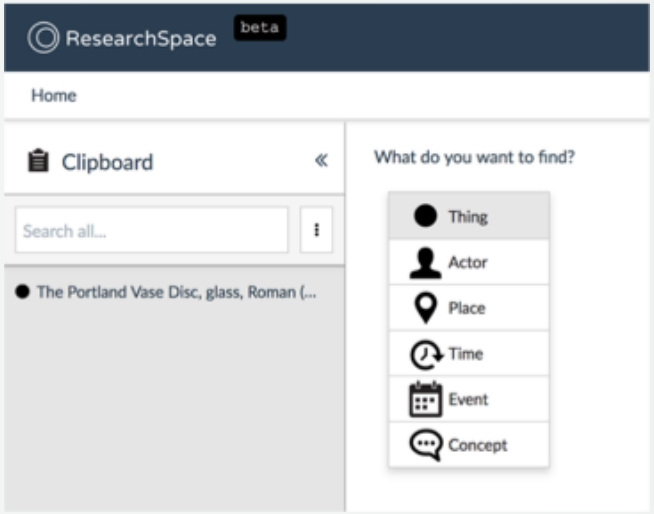
Planning	Mapping	Deploying	Use

- Deploy Platform
- Load
  - Ensure context
- Test
  - Formulate questions as queries
- Return to Mapping if Necessary

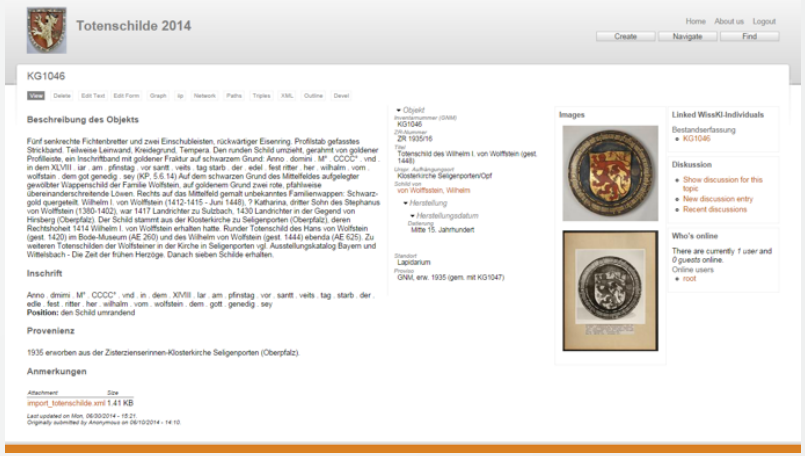
Use Overview

Planning	Mapping	Deploying	Use

- User Adoption
  - Query, Data Entry, Visualization, Navigation
- User Feedback
  - Researcher satisfaction
  - Return of data required
  - Input for platform improvement
  - Input for data improvement, expansion, Enrichment
  - Input for Query Improvement
- Return to Original Planning Phase

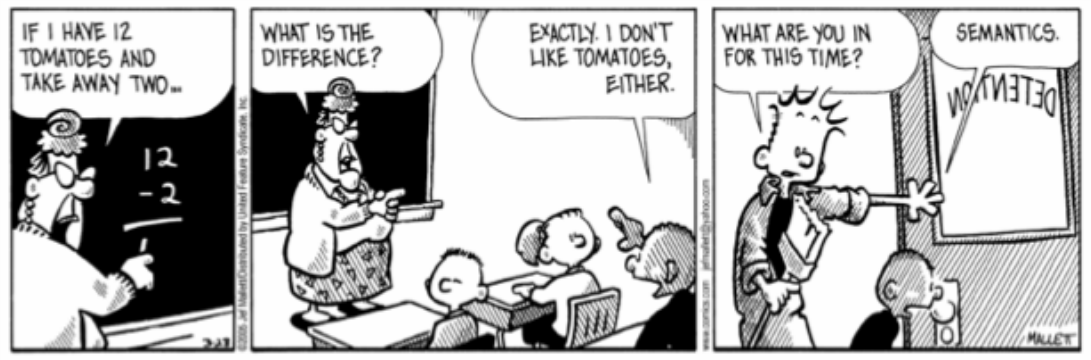


e.g. ResearchSpace Visual Query Mechanism



e.g. Wisski Semantic Data Display

Semantic Data and CRM, Why Again?



Research

- Ask and answer questions across ‘knowledge silos’ (in and across institutions and disciplines)
- Ask and answer more complex questions
- Greater knowledge discovery facility
- Rediscover and reanalyze ‘lost’ data

Information Management

- Improve data awareness / interoperability
- Support sustainability and reusability of data
- Software neutral exchange format

Semantic Data and CRM Adoption Learn More

CIDOC CRM Reference Materials

CIDOC CRM Specification
http://www.cidoc-crm.org/releases_table

Visual Charts
http://old.cidoc-crm.org/cidoc_graphical_representation_v_5_1/graphical_representation_5_0_1.html

Tutorials

One video

Many powerpoints
http://www.cidoc-crm.org/tutorialPage

Mailing list

Research Data Management Model

Parthenos Project
http://www.parthenos-project.eu/

(in progress)

X3ML Mapping Suite Tool

Open Access Service
https://www.ics.forth.gr/isl/3M

Source Code
http://github.com/isl/Mapping-Memory-Manager
https://github.com/isl/3MEditor
https://github.com/isl/x3ml

Semantic Platforms

Wiss-ki Project
http://wiss-ki.eu/

Research Space Project
http://www.researchspace.org/

End

Questions?

Dr. George Bruseker
ICS-FORTH
bruseker@ics.forth.gr