

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

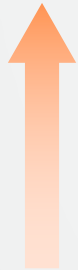
Refer
interpret
present



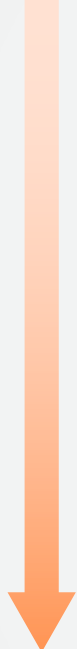
Search,
correlate,
integrate



discover
collect
aggregate
update



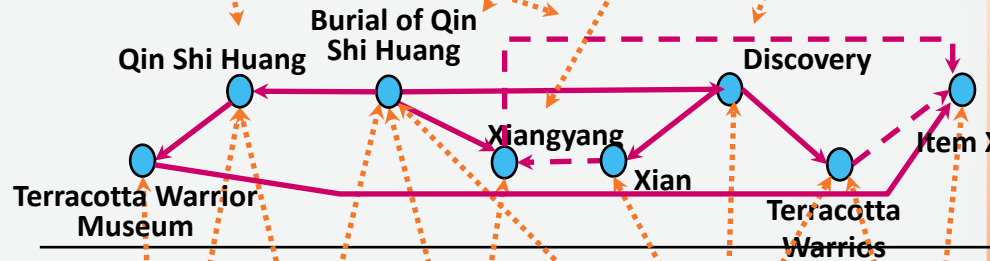
reexamine



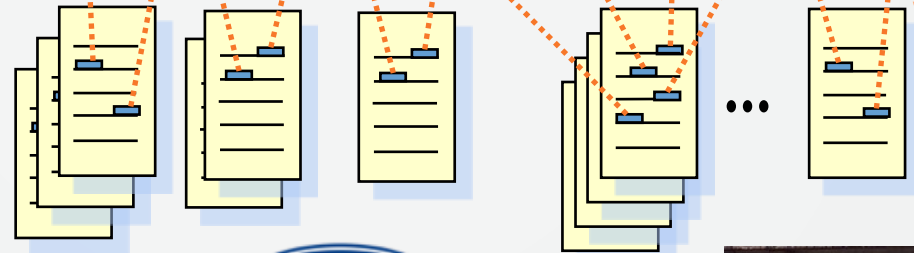
Publications
Stories
exhibitions



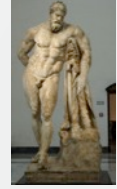
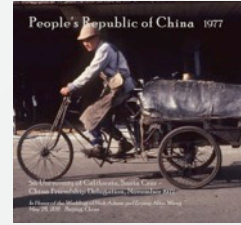
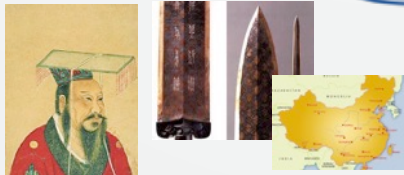
Layer of
"Latest
Knowledge"



"Evidence layer"
Things
Sources
Collections
Corpora



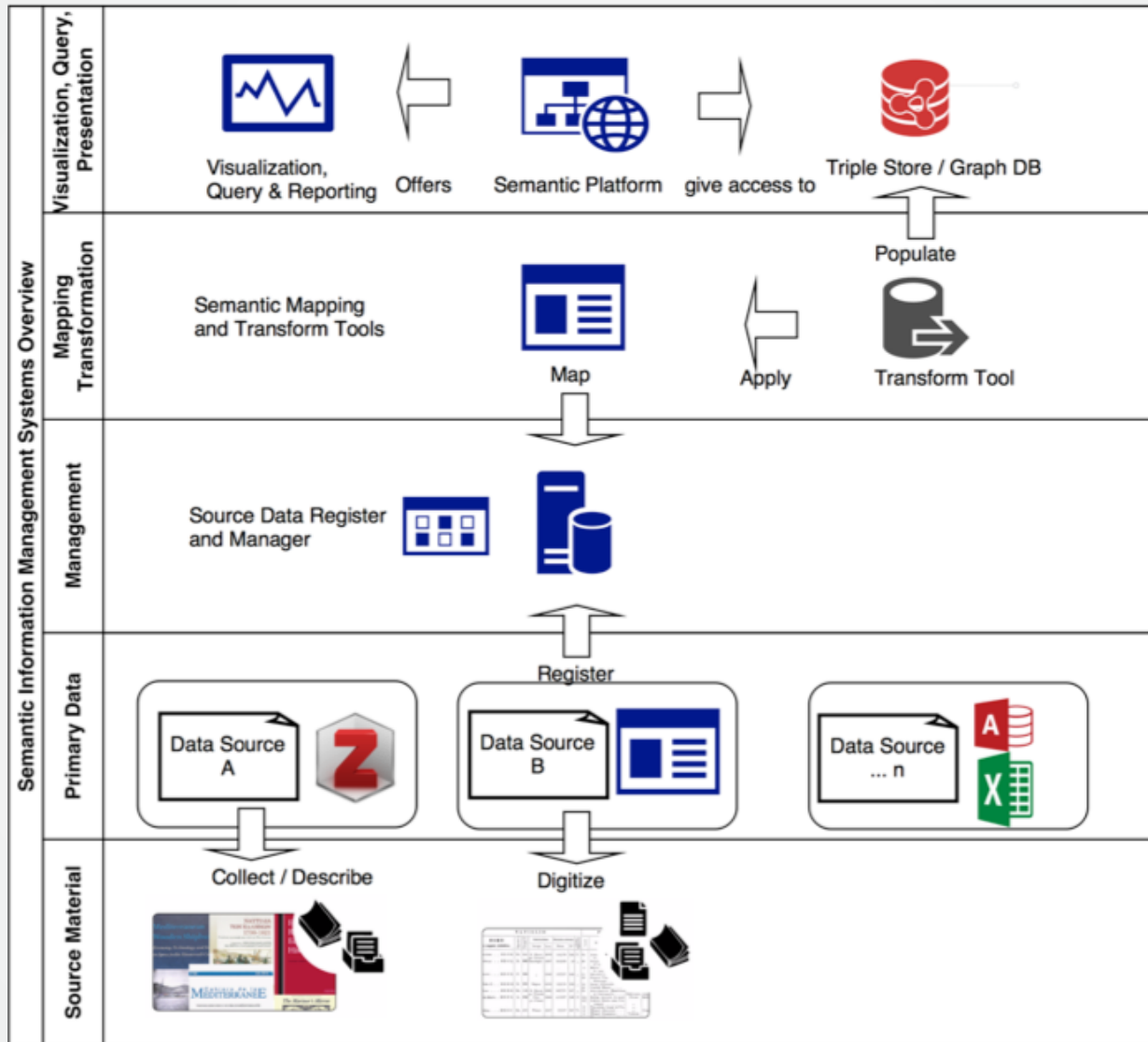
秦



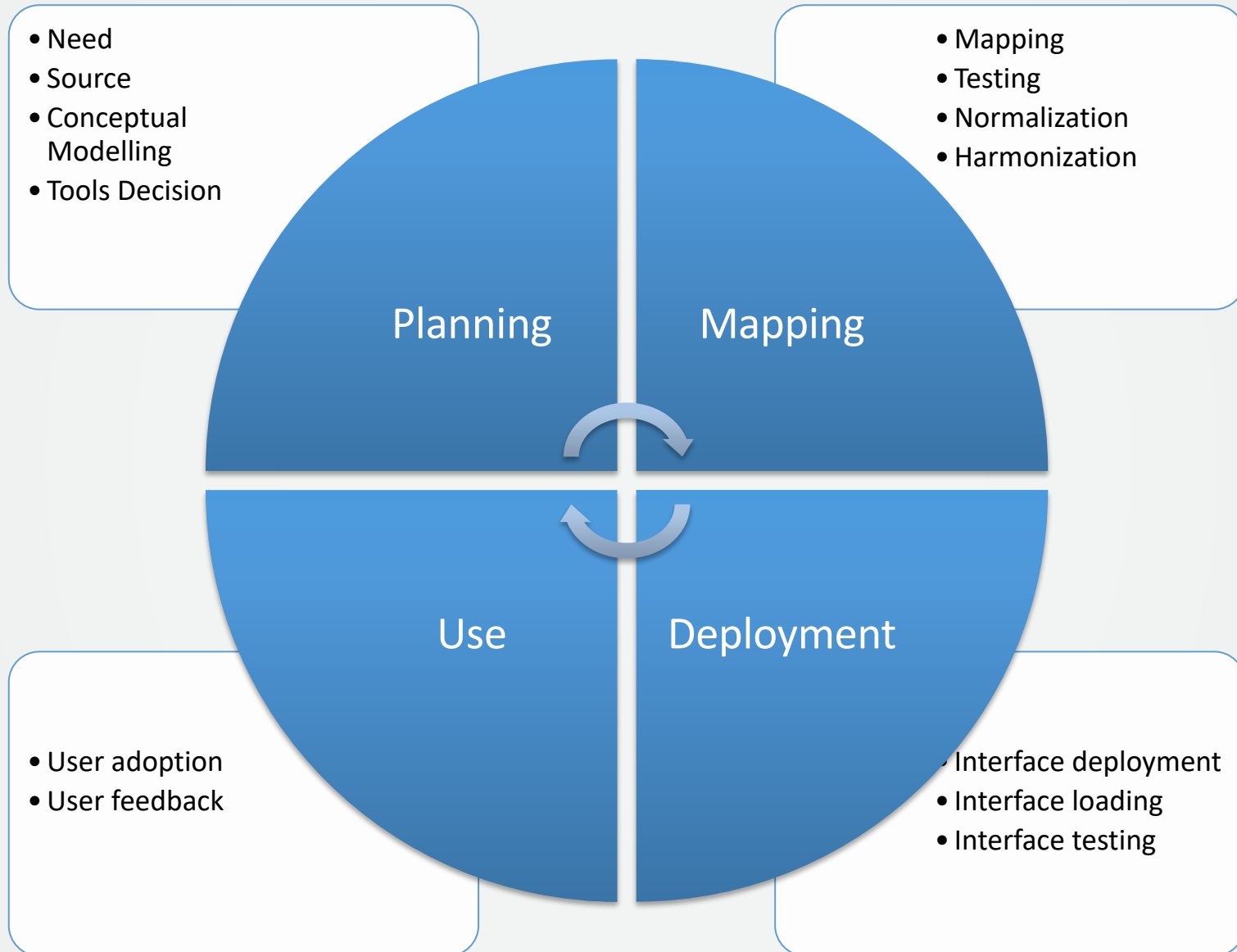
Curator,
Conservator,
Excavator



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

goals

- Common research or presentation

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

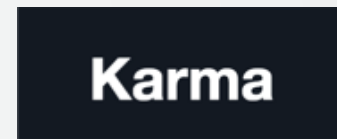
Planning	Mapping	Deploying	Use
Need	Sources	Modelling	Tools

Sources Register



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

Mapping Tools



Triple Stores / Graph DBs



Semantic Data Management Platforms



Mapping Overview

Planning

Mapping

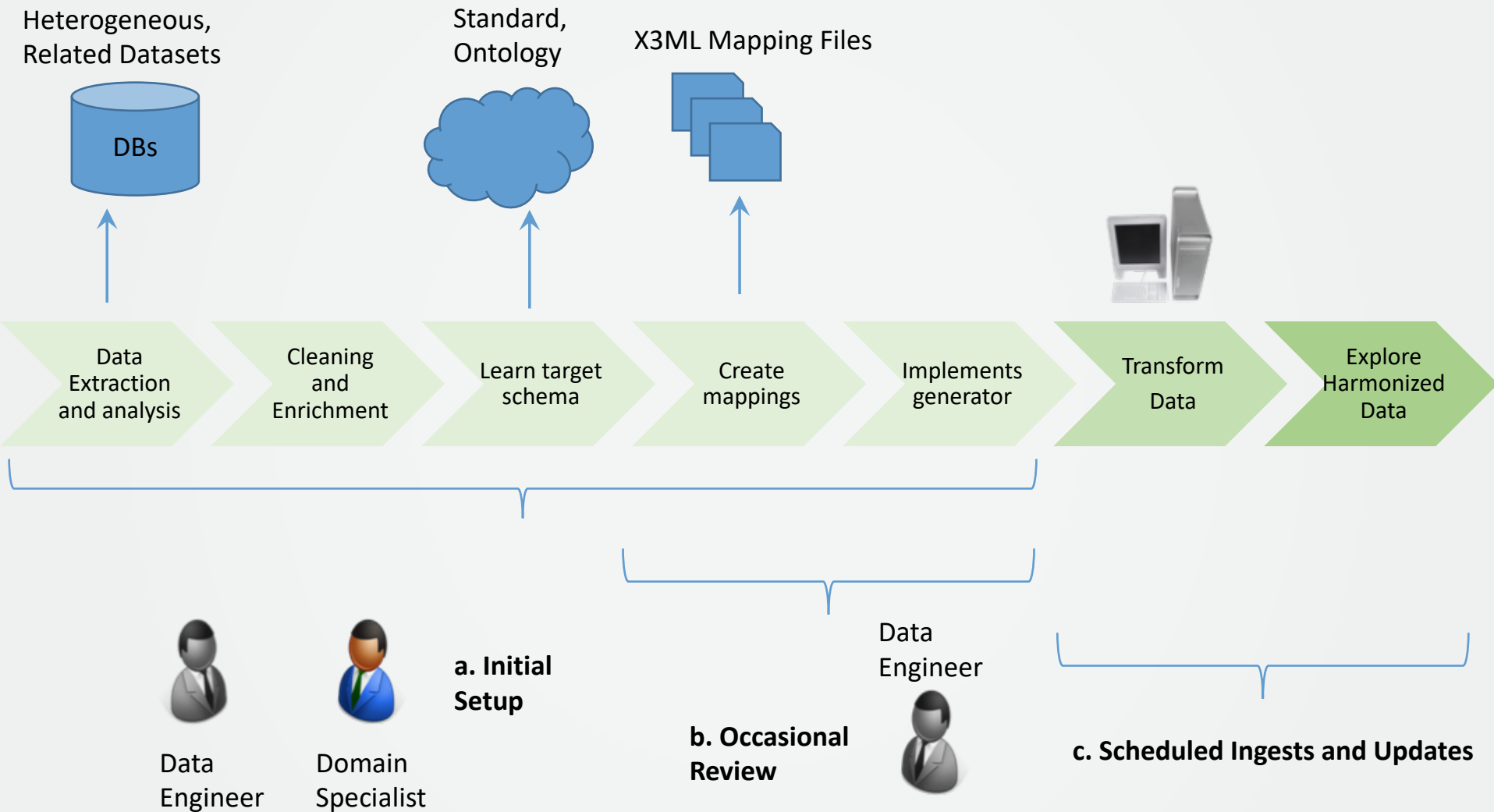
Deploying

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

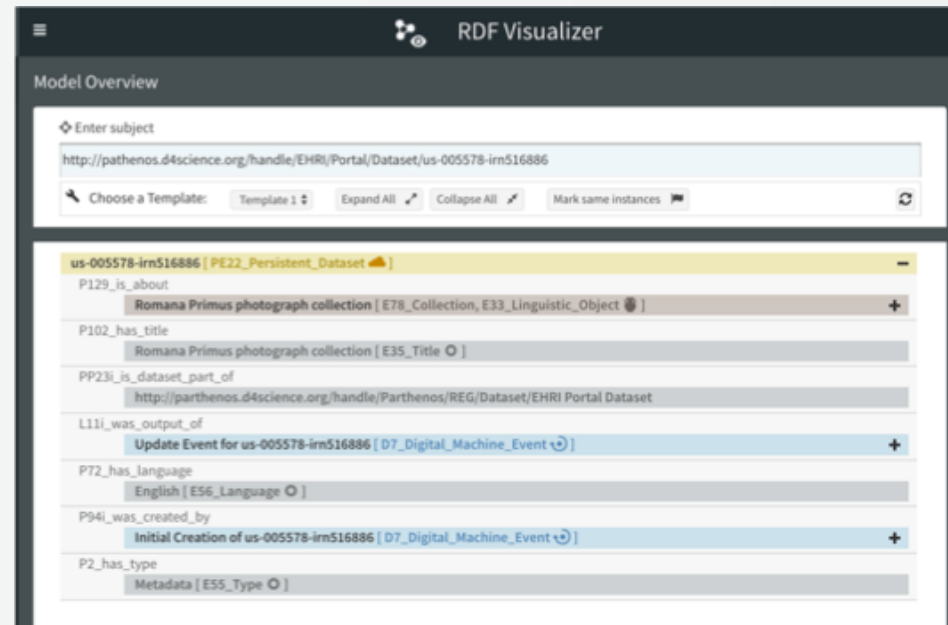
#	SOURCE	TARGET	CONSTANT EXPRESSION	IF RULE	COMMENTS
1	↓ /TOPOS	↓ E27_Site	[P2_has_type [E26_Type = "site"]]		
1.1	↓ KurzbeschreibungTopos	↓ PT_is_identified_by ↓ E41_Appellation	[P2_has_type [E26_Type = "official"]]		E41_Appellation URI can map to...
R	↓ KurzbeschreibungTopos	↓ rdf-schema:label			
1.2	↓ KurzbeschreibungTopos	↓ rdf-schema:label			
R	↓ KurzbeschreibungTopos	↓ rdf-schema:label			
1.3	↓ Lage	↓ P07_is_domain_of ↓ P03_has_note	[P51_has_type [E26_Type = "Topographic Position"]]		
R	↓ Lage	↓ P02_has_range ↓ rdf-schema:label			
1.4	↓ Ausdehnung	↓ P43_has_dimension ↓ E54_Dimension	[P91_has_unit [E26_Measurement_Unit = "square meters"]]		
R	↓ Ausdehnung	↓ P90_has_value ↓ rdf-schema:label	[P2_has_type [E26_Type = "surface area"]]		
1.5	↓ PS_ToposID == ↓ FS_ToposID	↓ P89_falls_within			
R	↓ Koordinate	↓ SP6_Declarative_Place			

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

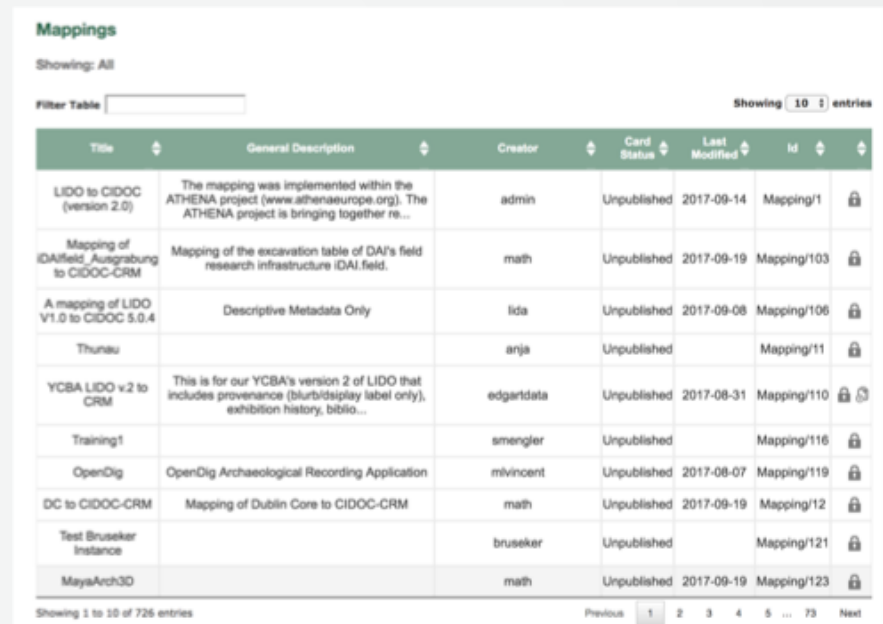
- Schema to Schema matching from a source to CRM leaves data values unchanged.
- Normalization of data values according to international standard thesauri ensures better interoperability.
- Normalization Operation should be documented.



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.



The screenshot shows a web interface titled "Mappings" with a search filter and a table of 10 entries. The table has columns for Title, General Description, Creator, Card Status, Last Modified, and Id. The entries include mappings from LIDO to CIDOC, IDAifield to CIDOC-CRM, YCBA LIDO v2 to CRM, and others.

Title	General Description	Creator	Card Status	Last Modified	Id
LIDO to CIDOC (version 2.0)	The mapping was implemented within the ATHENA project (www.athenaeurope.org). The ATHENA project is bringing together re...	admin	Unpublished	2017-09-14	Mapping/1
Mapping of IDAifield_Ausgrabung to CIDOC-CRM	Mapping of the excavation table of DAI's field research infrastructure IDAI field.	math	Unpublished	2017-09-19	Mapping/103
A mapping of LIDO V1.0 to CIDOC 5.0.4	Descriptive Metadata Only	lida	Unpublished	2017-09-08	Mapping/106
Thunau		anja	Unpublished		Mapping/11
YCBA LIDO v2 to CRM	This is for our YCBA's version 2 of LIDO that includes provenance (blurbsdisplay label only), exhibition history, biblio...	edgardata	Unpublished	2017-08-31	Mapping/110
Training1		smengler	Unpublished		Mapping/116
OpenDig	OpenDig Archaeological Recording Application	mivincet	Unpublished	2017-08-07	Mapping/119
DC to CIDOC-CRM	Mapping of Dublin Core to CIDOC-CRM	math	Unpublished	2017-09-19	Mapping/12
Test Bruseker Instance		bruseker	Unpublished		Mapping/121
MayaArch3D		math	Unpublished	2017-09-19	Mapping/123

Screenshot from X3ML tool suite;
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

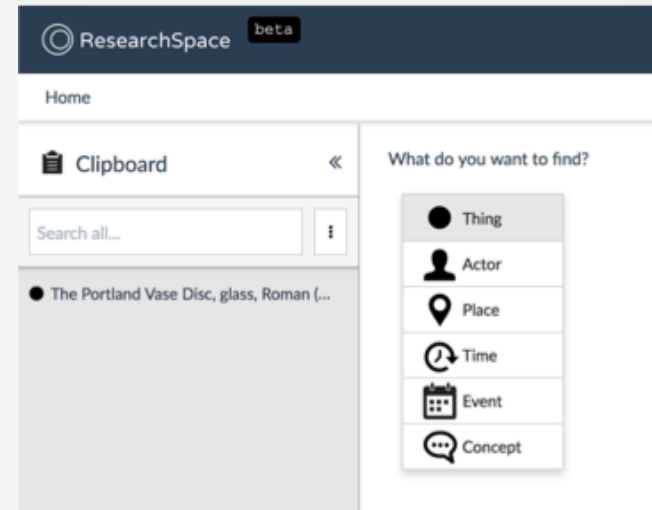
Planning

Mapping

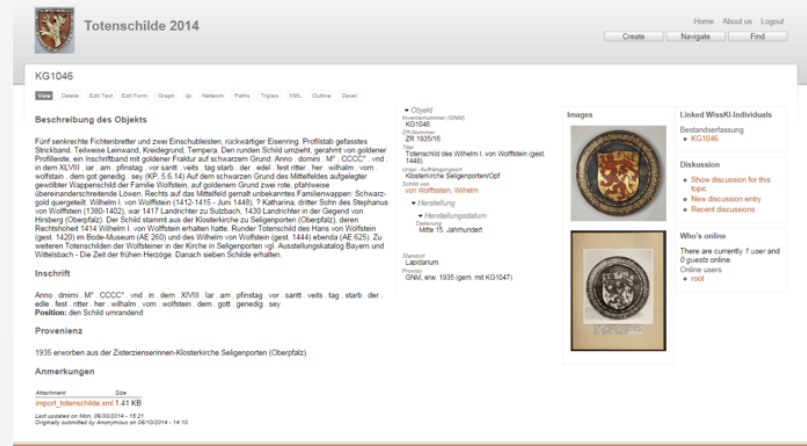
Deploying

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

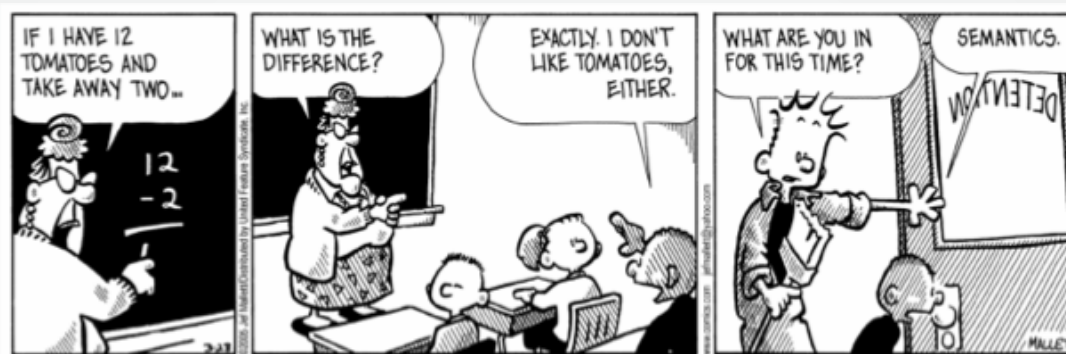


e.g. ResearchSpace Visual Query Mechanism



e.g. WisKi 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

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>

Research Data Management Model

Parthenos Project

<http://www.parthenos-project.eu/>

(in progress)

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