What WP7 can do for you

First ULICE Annual Meeting

Stockholm, September 3rd 2010

Deliverables to date

M 3	Dec 2009	
7.1	Federated security infrastructure; metadata services for recording terminology; model generation components for the creation of data capture interfaces	Bleddyn Jones

M 6	Mar 2010	
7.2	Selection, development and curation of metadata elements and models for the capture of inter-operable presentation, treatment and outcome data for large hadron-therapy	Bleddyn Jones

semantics

to use data *reliably* and *efficiently* outside the immediate context of its collection, we need a computable representation of its meaning
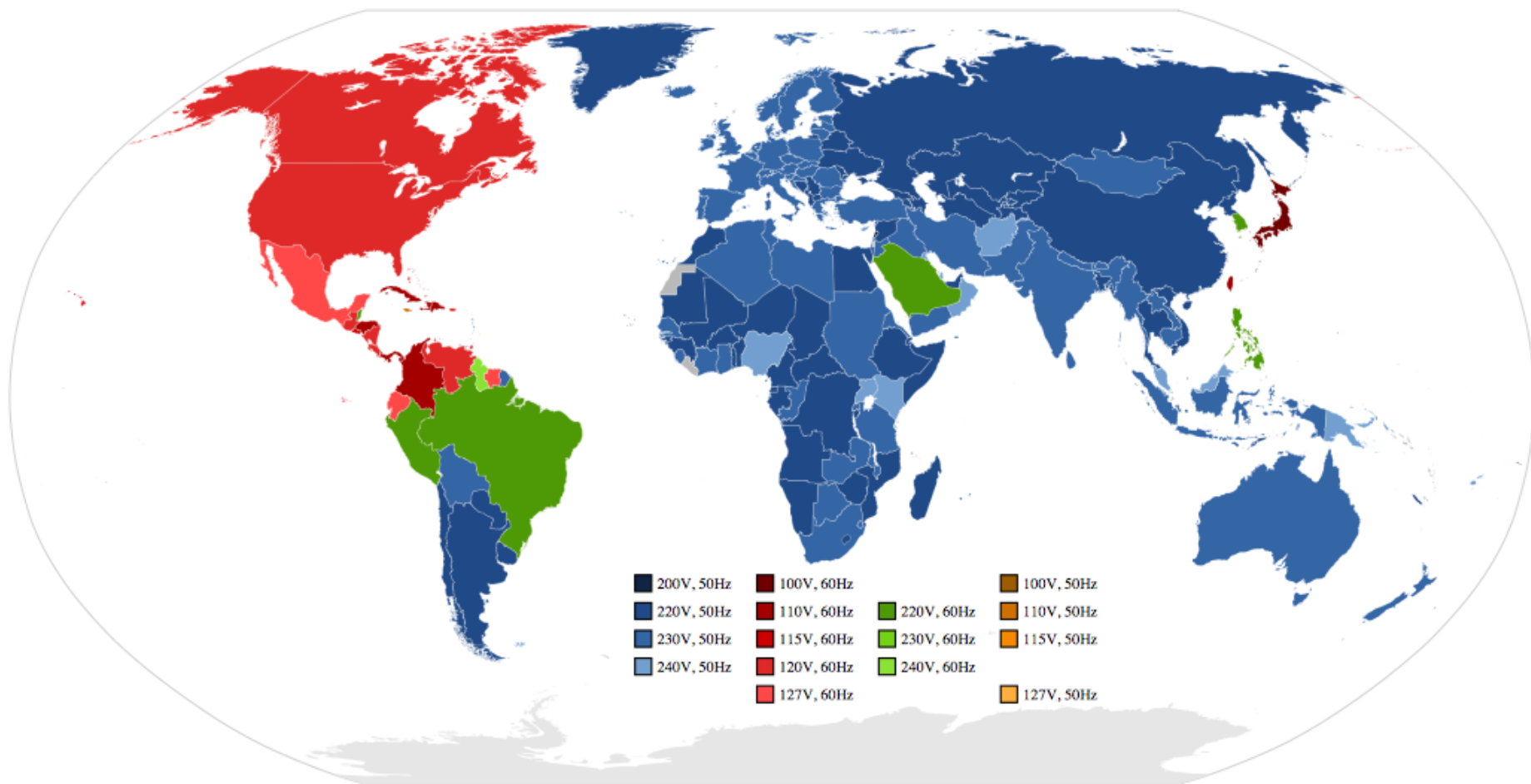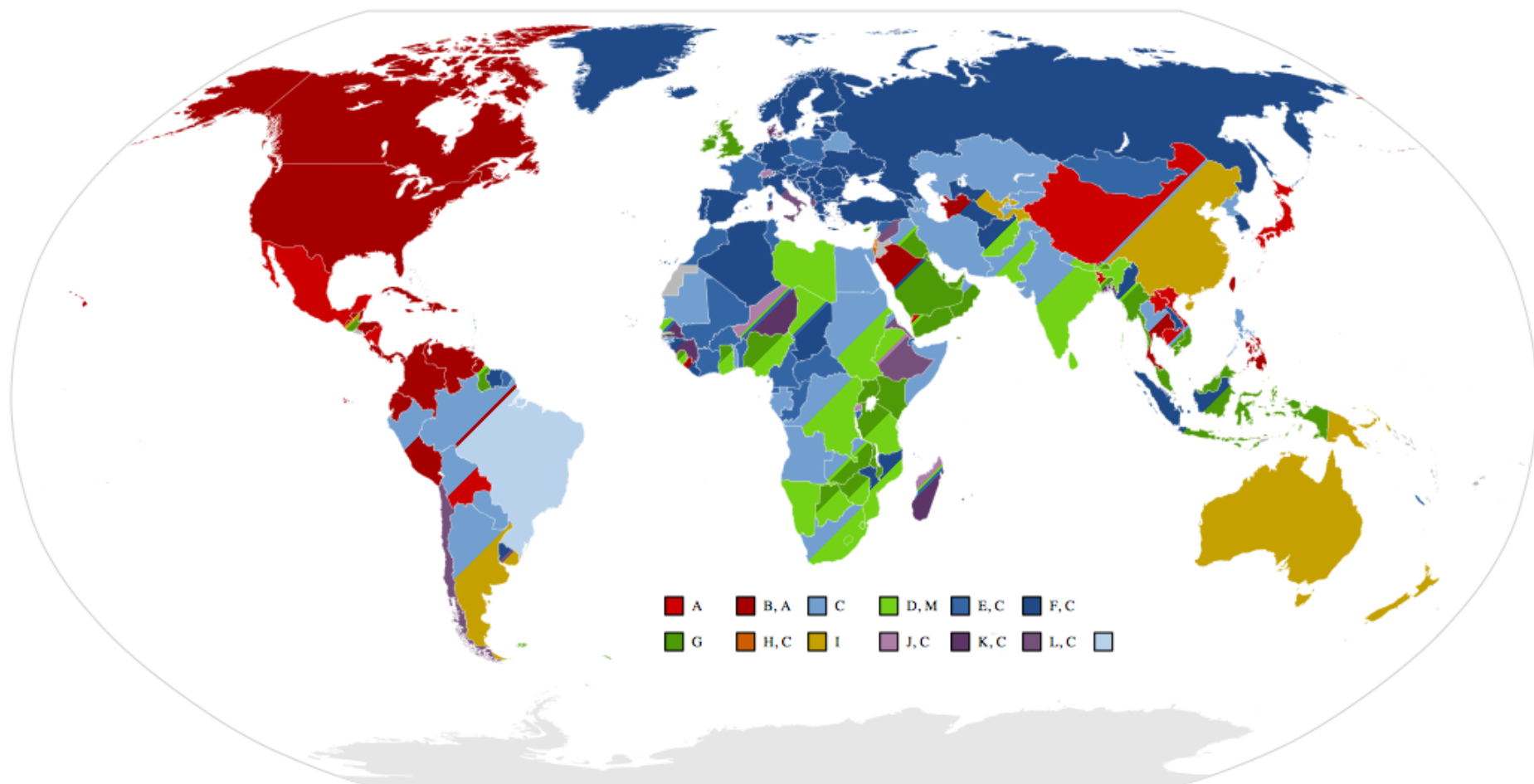
we need semantic metadata

semantic metadata

- value domain – range of possible values and their intended interpretations
- context
  - natural language definition of data item
  - PDF of form in which it appears
  - documentation for workflow or protocol
  - *schema* for form, or database, or workflow

semantic interoperability

- semantics are local, temporary, and specific to purpose
- we need to support multiple standards, multiple versions of standards, because
  - this is *research* and we expect *change*
  - we are working in multiple contexts

■ 200V, 50Hz	■ 100V, 60Hz		■ 100V, 50Hz
■ 220V, 50Hz	■ 110V, 60Hz	■ 220V, 60Hz	■ 110V, 50Hz
■ 230V, 50Hz	■ 115V, 60Hz	■ 230V, 60Hz	■ 115V, 50Hz
■ 240V, 50Hz	■ 120V, 60Hz	■ 240V, 60Hz	■ 127V, 50Hz
	■ 127V, 60Hz		

A	B, A	C	D, M	E, C	F, C
G	H, C	I	J, C	K, C	L, C

WP7

- metadata registry
  - record and publish *your* value domains and data definitions
  - add *models* of forms, schemas, protocols, proposed standards
  - relate your data, your models, your work to the work of others

WP7

- metadata services
  - annotate existing data sets, forms, protocols, models, with links to published definitions
  - instantiate a *metamodel* for clinical trials to produce a model of your own trial
  - generate forms *and trial management services* from your own model

open for business

the technology

- – is ready now, works on standard platforms, and is entirely open (BSD license)

- – will be maintained, supported, and enhanced to address user requirements

- – can be used *locally* and federated; it is also interoperable with the US NCI provision

MRC Medical Research Council

Leading science for better health

cancergrid metadata registry

Data Element: ECOG Performance Status

as XML

Administered Item - Preferred Name: ECOG Performance Status

Administered Item Identifier	GB-CANCERGRID-D1E682341-0.21
Registration Status	Recorded
Registered By	Dr. Peter Macallum Project Manager, CancerGrid 01223 765348

Designation

Context	**Clinical Trials**
Name	ECOG Performance Status (preferred designation)
Definition	Used to assess how a disease is progressing, how the disease affects the daily living abilities of the patient, and determine appropriate treatment and prognosis.
Language	GB-eng
Definition Source Reference	

Context	**Clinical Trials**
Name	Summario dei dati per ECOG
Definition	Definizione Usato per valutare come la malattia evolve, come la malattia influenza le attivita' di vita quotidiana del paziente e per determinare trattamento e prognosi adeguati
Language	Lita

Type	enumerated
Full Record	**ECOG Performance Status**

Summary	permissible value	value meaning	reference
	0	Normal	**urn:lsid:ncicb.nci.nih.gov:nci-thesaurus:C19998**
	1	Symptomatic: fully ambulatory	**urn:lsid:ncicb.nci.nih.gov:nci-thesaurus:C19999**
	2	Symptomatic: in bed less than 50% of the time	**urn:lsid:ncicb.nci.nih.gov:nci-thesaurus:C17846**
	3	Symptomatic: in bed greater than 50% of the time	**urn:lsid:ncicb.nci.nih.gov:nci-thesaurus:C17847**
	4	100% bedridden	**urn:lsid:ncicb.nci.nih.gov:nci-thesaurus:C17848**
	5	dead	**urn:lsid:ncicb.nci.nih.gov:nci-thesaurus:C20000**

Related Data Elements

asserted by this data element	id	preferred name	how related	status
	GB-CANCERGRID-D1E573861-0.19	ECOG performance status is 0, 1 or 2	narrowerTerm	Recorded
	US-NCICB-88-5.1		equivalentTo	

asserted by other data elements	id	preferred name	how related	status
	GB-CANCERGRID-D1E5738610.19	ECOG performance status is 0, 1 or 2	broaderTerm	Recorded

Reference Documents

id	language	title	type
KNCSJG1HE	eng	Breast Cancer Phase III Reporting standards.pdf	commissioning
ecog_amjclinoncol	eng	Oken, M.M., Creech, R.H., Tormey, D.C., Horton, J., Davis, T.E., McFadden, E.T., Carbone, P.P.: Toxicity And Response Criteria Of The Eastern Cooperative Oncology Group. Am J Clin Oncol 5:649-655, 1982.	Journal paper
ecog_perf_status	eng	ECOG Performance Status	web page
tango_rand_form	eng	Tango Clinic Randomisation Form version 3	Example usage

	GB-CANCERGRID-D1E5738610.19	ECOG performance status is 0, 1 or 2	broaderTerm Recorded

Reference Documents

id	language	title	type
KNCSJG1HE	eng	Breast Cancer Phase III Reporting standards.pdf	commissioning
ecog_amjclinoncol	eng	Oken, M.M., Creech, R.H., Tormey, D.C., Horton, J., Davis, T.E., McFadden, E.T., Carbone, P.P.: Toxicity And Response Criteria Of The Eastern Cooperative Oncology Group. Am J Clin Oncol 5:649-655, 1982.	Journal paper
ecog_perf_status	eng	ECOG Performance Status	web page
tango_rand_form	eng	Tango Clinic Randomisation Form version 3	Example usage

Conceptual Framework

Expresses data element concept (full record)	ECOG Performance Status data element concept	
Concept or concept expression	Study Subject Performance Status	urn:lsid:ncicb.nci.nih.gov:nci-thesaurus:C25400
Property measured	Code	urn:lsid:ncicb.nci.nih.gov:nci-thesaurus:C25162
Typed by representation class	Code	
Representation class qualifier	unqualified	

Administration

Administrative Status *explain*	scheduledForReview
Administered By	Peter MacCallum Project Manager, CancerGrid
Creation On	2005-03-15
Effective From	2006-06-15
Last Changed On	2008-04-22
Effective until	
Submitted By	Sui L lin Researcher, CancerGrid
Explanatory Comments	Auto generated from Protege

Home | Insert | Page Layout | Formulas | Data | Review | View

From Access | From Web | From Text | From Other Sources | Existing Connections | Refresh All | Connections | Properties | Edit Links | Sort | Filter | Clear | Reapply | Advanced | Text to Columns | Remove Duplicates | Data Validation | Consolidate | What-If Analysis | Group | Ungroup | Subtotal | Show Detail | Hide Detail

Get External Data | Connections | Sort & Filter | Data Tools | Outline

C5

	A	B	C	D	E	F	G	H
1	Patient Trial Number (CRUK Trials Unit, Birmingham)	Study Subject Full Name	ECOG Performance Status					
2	1	John Doe	0 : (Normal)					
3	2	James Harrison	2 : (Symptomatic: in bed less than 50% of the time)					
4	3	Peter McBride	1 : (Symptomatic: fully ambulatory)					
5	4	Matt Botany						
6	5	Lionel Marks	0 : (Normal)					
7			1 : (Symptomatic: fully ambulatory)					
8			2 : (Symptomatic: in bed less than 50% of the tin					
9			3 : (Symptomatic: in bed greater than 50% of th					
10			4 : (100% bedridden)					
11			5 : (Dead)					

	E	F	G
	Preferred Name	Definition	Properties/Value Domain
21	ECOG Performance Status	Used to assess how a disease is progressing, how the disease affects the daily living abilities of the patient, and determine appropriate treatment and prognosis.	0 : (Normal), 1 : (Symptomatic: fully ambulatory 2 : (Symptomatic: in bed less than 3 : (Symptomatic: in bed greater th time), 4 : (100% bedridden), 5 : (Dead)
	GB-CANCERGRID-0A296A56A-0.14	Study Subject Full Name	Full name of patient as normally written by subject. Order of given and family names is that normally written by the subject.	xs:string
	GB-CANCERGRID-D1E592581-0.16	Patient Trial Number (CRUK Trials Unit, Birmingham)	A sequential, integer trial identifier is assigned to each eligible, recruited patient at registration. Occasionally, numeric ranges within the domain of allowed trial numbers are assigned to offer additional semantics: for example, in several trials, trials numbers of 10,000 or greater were issued when automatic treatment allocation was unavailable and the treatment was allocated manually by randomisation room staff	xs:positiveInteger

File   Edit   View   Project   Diagram   Element   Tools   Add-Ins   Settings   Window   Help

Logical Diagram: "Class Model"   created: 19/11/2005 modified: 10/01/2006 100% 827 x 1169

«enumeration»
Enum_ECOG_Performance_Status

«enum»
0
1
2
3
4
5

ECOG_Performance_Status

+   definition: string = Used to assess ... {readOnly}
+   id: string = GB-CANCERGRID-D... {readOnly}

CancerGrid Query Service

Free Text   Classification

Controls

cgMDR-CG1

ecog

Query completed.   Search

Results

ECOG Performance Status
ECOG performance status is 0, 1 or 2

<<   >>   Use

Details

Definition   Props/Values

Code	Meaning
0	Normal
1	Symptomatic: fully ambulatory
2	Symptomatic: in bed less than 50% of the time
3	Symptomatic: in bed greater than 50% of the time
4	100% bedridden
5	unknown death dead

Start Page   *Class Model

System   Output

Default Style

Ready   CAP NUM SCRL

start   bin   Tomcat   test2 - EA - 30 Day Trial   CancerGrid Query Se...   00:04

openMDR

THE OHIO STATE UNIVERSITY Medical Center

- Federated semantic metadata management utilizing and enhancing UK CancerGrid cgMDR.

Data Elements Services
NCI · OSU · WashU · UCSF

Domain Model
publish model
local cgMDR

Ontologies/Terminologies Services
NCI · OSU · WashU · UCSF
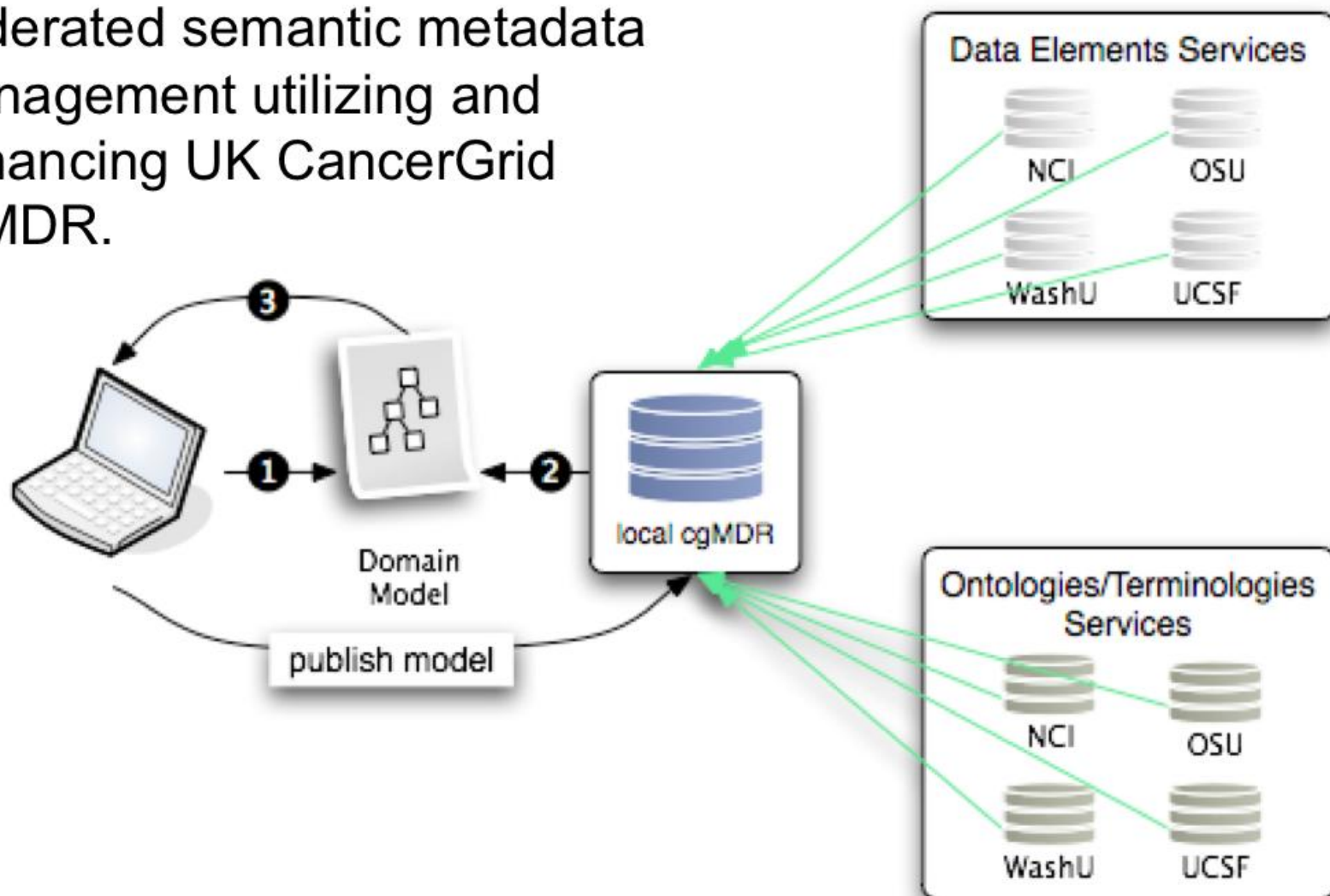
Commons
Background
Repository
Forum
Congress

Training
Presentations
Center for Cancer
Systems Biology

Sage
Careers
Directors
Partners
News
FAQs

Research
Case Studies
Publications
Resources
Tools

Sage
BIONETWORKS
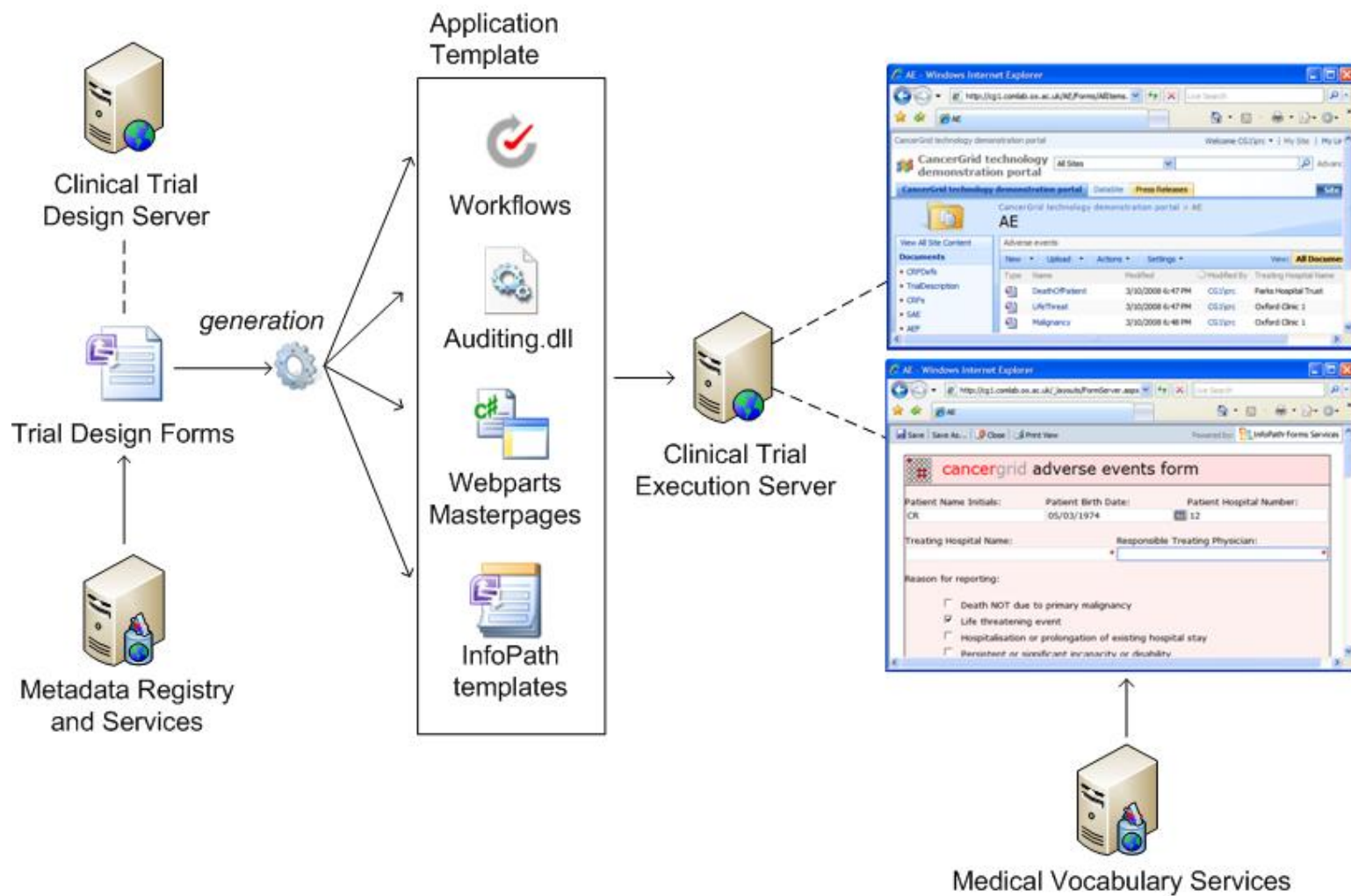
File Edit View Insert Format Tools Table CancerGrid Help

Type a question for help

Metadata   Errors

Case Report Form Design

cancergrid
trial designer

Trial event: eligibility
Intervention codes:
☐ Insert intervention code
Form control name: Eligibility
Namespace prefix: CCVTM
Namespace: http://www.cancergrid.org/CCVTM/crf/2009-01-14/

Add and select controls in this table then edit the details for the control below.

Eligibility	Eligibility
InclusionCriteria	Inclusion Criteria
ExclusionCriteria	Exclusion Criteria
ParentConsent	the parent or guardian of the child to be enrolled into the study is willing and able...
ChildAgeRequirement	child aged 3 years 4 months and above

☐ Insert control

General control details
Control name: InclusionCriteria
Title: Inclusion Criteria
Tool tip:
○ Click here to insert expression describing when to disable the control
○ Click here to insert expression describing when control data is invalid

Section
Contained Control Names
ParentConsent
ChildAgeRequirement
☐ Insert item
Protocol:
Participants eligible to be enrolled into the study

The following set of concepts describe the context of this section:

Concept

Form template's location: http://cg1.comlab.ox.ac.uk

Query Service Control

Free Text | Classification

Controls
cgMDR
child
Query completed.   Search

Results
Common Toxicity Criteria Adverse
child aged 3 years 4 months and a
parent or guardian willing and able

<   >   Use

Details
Definition | Props/Values

please complete definition...

🔲 ACWY

View All Site Content

Documents
- Shared Documents
- Centres
- Researchers
- Case Report Forms
- SAE
- Screening Failures

Lists
- Calendar
- Tasks
- KPI List

Discussions
- Team Discussion

Sites

People and Groups
- 🔲 **Recycle Bin**

Oxford Vaccine Group > ACWY

Understanding the Immune Response to Meningitis Vaccines

Files and records for the ACWY study

Shared Documents ▼

Type	Name	Modified	Modified By
📄	4-4-smpc-acwy-vax-vaccine	25/08/2009 11:13	CG1\mahshi
📄	Adult ACWY CRF 19 06 08 clean	02/07/2009 17:12	CG1\sharris
📄	Adult MenACWY 19 June 2008	02/07/2009 17:14	CG1\sharris
📄	MenACWY IB Edition 12 23 FEB 09	14/07/2009 09:37	CG1\mahshi
📄	Participant Contact Sheet	02/07/2009 17:14	CG1\sharris

※ Add new document

Key Performance Indicators ▼

Indicator	Status
Total Recruitment	🔴
Protocol Violations	🟢
Serious Adverse Events	🟢
Recruitment Rate	🔴

Case Report Forms ▼

New ▼ | Upload ▼ | Actions ▼ | Settings ▼

Name	Participant No	DOB	Initials	Visit 1	Visit 2	Visit 3	Visit 4	Visit 5	Withdrawn	Violations	AE Count
	Count = 6			Count = 6	Count = 4	Count = 0	Count = 0	Count = 0	Count = 0	Sum =	Sum = 1
001	001			20/08/2009	27/08/2009						1
002	002			20/08/2009	27/08/2009						
003	003			25/08/2009	01/09/2009						
004	004			27/08/2009	02/09/2009						
005	005			01/09/2009							
006	006			01/09/2009							

SAE ▼

New ▼ | Upload ▼ | Actions ▼ | Settings ▼

Name	Centre	Researcher	Participant No	DOB	Initials	Onset	Resolved	FU Planned	FU Completed

To create a new item, click "New" or "Upload" above. There are no items to show in this view of the "SAE" document library.

METABRIC

Heterogeneous clinical and laboratory data produced by different database systems

Addenbrooke's
Data

Guy's
Data

Manitoba
Data

Nottingham
Data

Vancouver
Data

Pathologist encodes domain knowledge relating different hospital data collections

- LN-, no chemo
- ER+, LN+, hormone, no chemo
- ER-, LN+, chemo, no taxanes
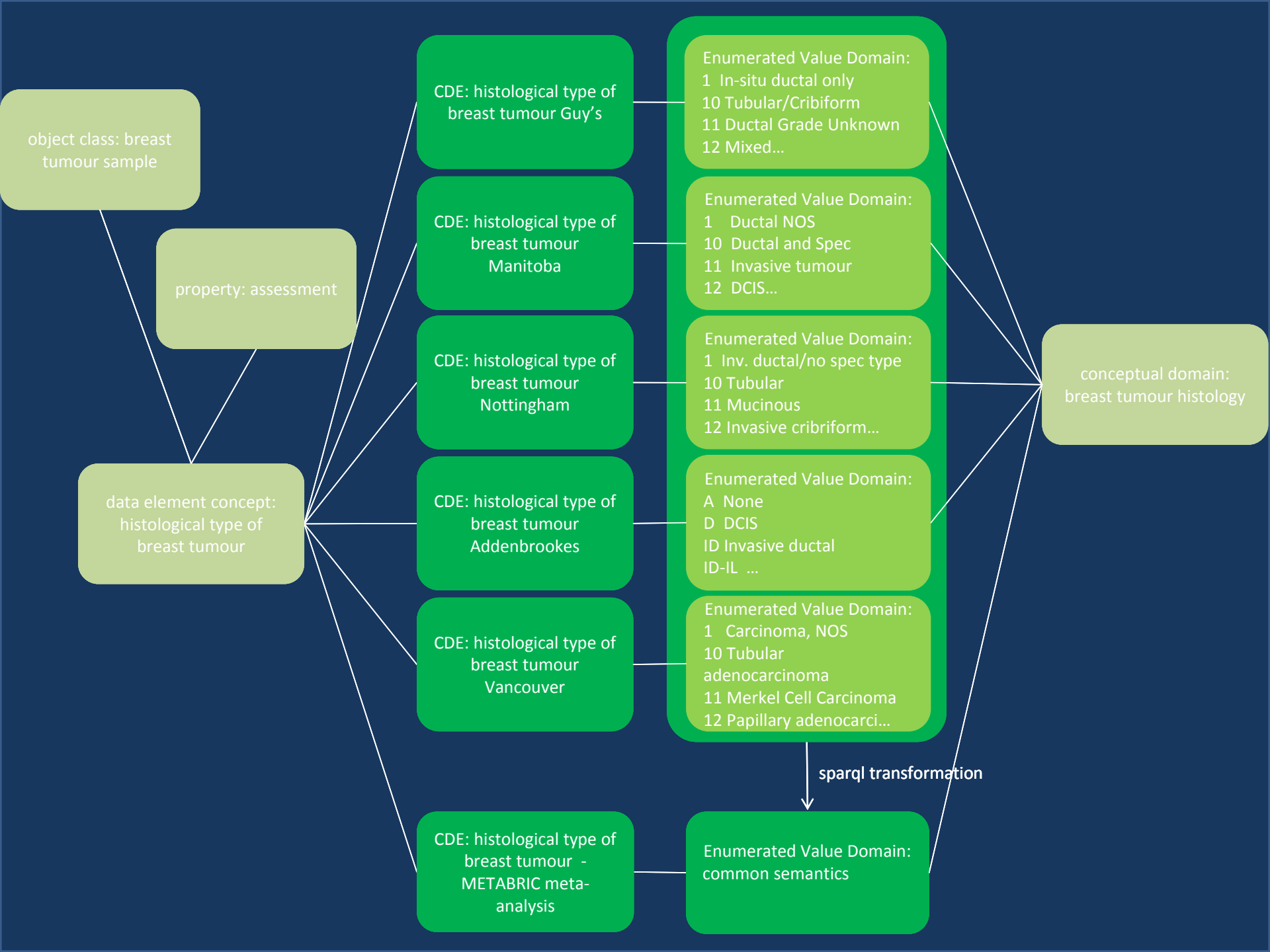- ...

Metadata databases

CancerGrid SQIV tool suite performs data integration of clinical and laboratory data using inference rules

Homogeneous data (2000 breast cancer tumours)

Data mining of molecular signatures and taxonomies

steve.harris@comlab.ox.ac.uk

jim.davies@comlab.ox.ac.uk