

Practical and informatics challenges of using data from the electronic health record: From ePCRn to TRANSFoRm

Professor Brendan Delaney
Kings College London

www.transformproject.eu

History

ePCRn

- Funder NIH (US) 2005-8 Pilot for NIH Roadmap
- University of Minnesota, University of Birmingham, UCSF.
- Additional funds 2008-9 from NIHR NSPCR (Bham)

TRANSFoRm

- Funder EU FP7 ICT, €7m integrated project 2010-15
- 16 partner universities in 9 EU countries led by KCL
- Builds on and links to ePCRn

Aims of ePCRn

Data linkage between eHRs and clinical research

- To identify eligible subjects

A standards-based system for generating web-based data capture for RCTs

- To control data items via a meta-data repository
- For re-use of data elements
- To generate forms and databases from the meta data

ePCRN brings the technologies of the NCI cancer Bioinformatics Grid (CaBIG) into the Primary Care setting

- **Extends the BRIDG data model**
- **Integrates the Enterprise Vocabulary Service**
- **Enables the creation and reuse of data elements and forms via the Data Standards Repository**

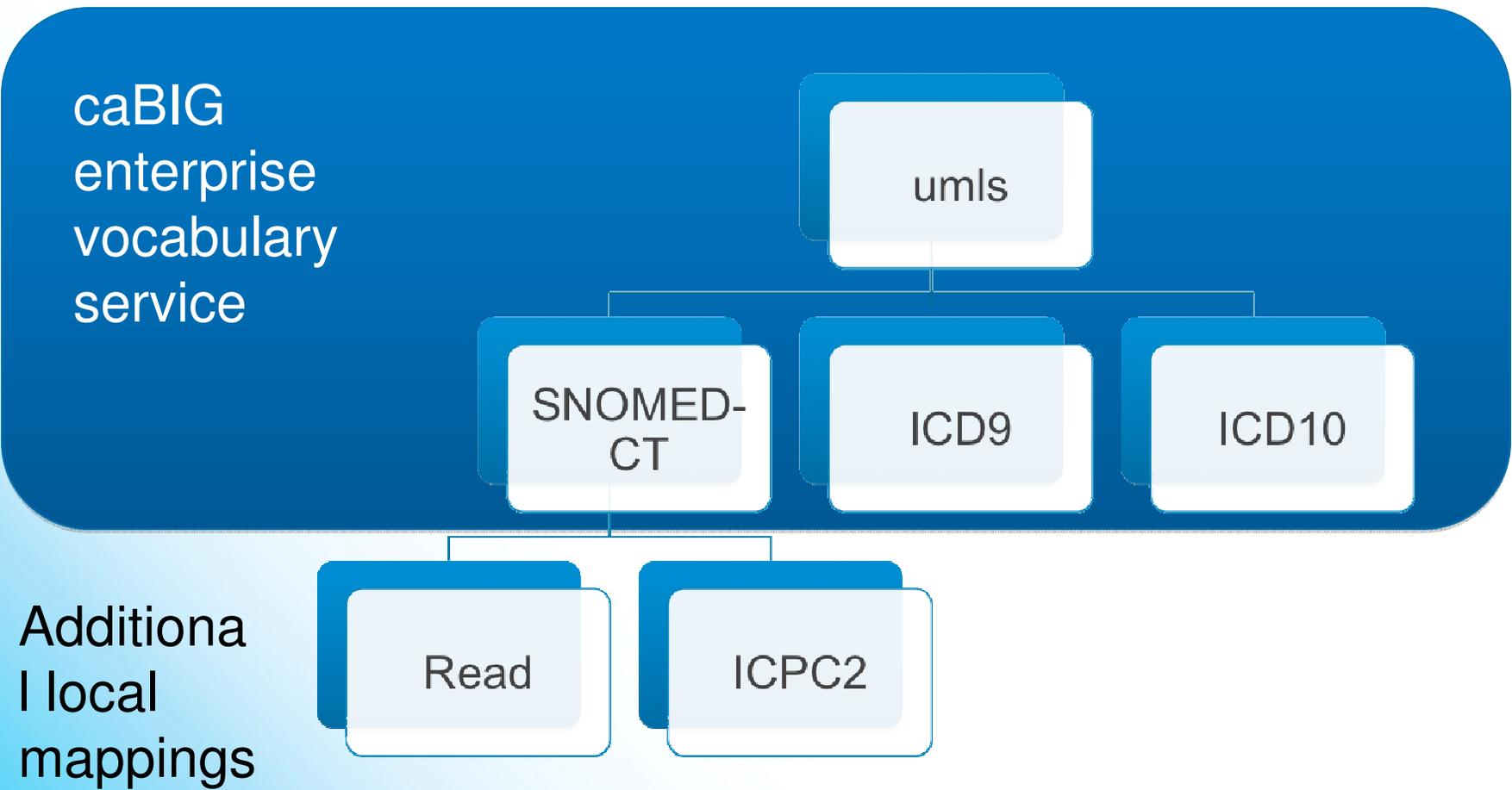
Confidentiality and data security

- **Confidentiality - for subjects not having consented to use of their data for research**
 - Search reports only 'counts' - NO DATA is extracted
 - Subjects are flagged locally
- **Security**
 - OGSA-DAI & GTK4(certificates and authorisation)

ePCRN: Recruitment

- 1. Capture protocol definition**
- 2. Drill down to capture eligibility criteria via 'semantic workbench'**
- 3. Browse CaBIG Enterprise Vocabulary Service**
- 4. Construct query in selected vocabulary**
- 5. Authorise and run**

Controlled vocabulary



Define eligibility criteria

The screenshot displays the ePCRN Study Designer v1.0 (Alpha) interface. The main window is titled "ePCRN Study Designer v1.0 (Alpha)" and contains a menu bar (File, Edit, Tools, Window, Help) and a toolbar. The active tab is "*Eligibility Criteria", with sub-tabs for "Inclusion Criteria", "Exclusion Criteria", and "EVS UI".

The "Define Age and Gender" section includes:

- AGE: \geq 20 years and \leq 60 years. An optional second age range is indicated by a right-pointing arrow.
- Gender: includes Male and Female.

The "Define Included Clinical Problems" section includes:

- Problem 0: includes all, Status: Active, Problem: Diabetes, None, 0 days. Includes search buttons (+, -, *) and a "Search EVS" button.

The "Eligibility Criteria Text" section displays the generated text: **Include Male and Female** Patients with age \geq 20 AND \leq 60.

The "Details" panel on the right contains a table with columns "Problem Name", "Status", and "Time F". Below it is a "Vocabulary and Codes" section with a table with columns "Vocabulary" and "Code".

Define the problem semantically

The screenshot displays the ePCRN Study Designer v1.0 (Alpha) interface. The main window is titled "ePCRN Study Designer v1.0 (Alpha)" and has a menu bar with "File", "Edit", "Tools", "Window", and "Help". Below the menu bar is a toolbar with various icons. The main area is divided into several sections:

- Enter Term:** A search box containing "Diabetes" and a "Search" button. Below it are radio buttons for "Thes" (selected), "MThes", and "By Code", along with a dropdown menu set to "ALL" and a numeric input set to "10".
- Found Matching Concepts:** A list box containing "Diabetes_Inspidus", "Diabetes_Management_Nurse_Specialist", "Diabetes_Mellitus" (highlighted), and "Diabetes_in_Pregnancy". A large red arrow points to "Diabetes_Mellitus".
- Identifiers:** A section with "NCI Code:" set to "C2985" and "Preferred:" set to "Diabetes Mellitus". Below are buttons for "Add this concept to EC", "Replace Selected Concept", and "Cancel".
- Information about this Concept:** A table with columns "Property Name" and "Value".

Property Name	Value
Synonym	diabetes mellitus
Preferred_Name	Diabetes Mellitus
Synonym	Diabetes
DEFINITION	<def-source>NCI</def-source><d...
- Super concepts:** A list box containing "Endocrine_Pancreas_Disorder" and "Glucose_Metabolism_Disorder".
- Sub concepts:** A list box containing "Diabetes_in_Pregnancy", "Insulin_Dependent_Diabetes", "Lipoatrophic_Diabetes_Mellitu", and "Non-Insulin_Dependent_Diabe".
- Details:** A section with "Available Codes" for "C0011849". It has radio buttons for "Common Codes" (selected) and "All Codes", and a "Select All" button. Below is a table of available codes.

Code Abbrv	Term	Code Value
<input checked="" type="checkbox"/> UMLS_CUI	Diabetes Mellitus	C0011849
<input checked="" type="checkbox"/> SNOMEDCT_2006_07_31	Diabetes mellitus	73211009
<input checked="" type="checkbox"/> SNOMEDCT_2006_07_31	Diabetes mellitus (disorder)	73211009
<input checked="" type="checkbox"/> SNOMEDCT_2006_07_31	DM - Diabetes mellitus	73211009
<input checked="" type="checkbox"/> ICD9CM 2007	Diabetes mellitus	250
- Messages:** A log area showing messages like "thread completed: Super Concepts loaded....", "thread started: Super Concepts loading....", etc.

Count eligible community subjects

EVSGUI V0.58

File Tools Window

Inclusion Criteria Exclusion Criteria EVS UI **EC Results** Query Builder

Query Text

by Code by Name LT Absent VS Absent

```

SELECT count(*) as count
FROM ccr_xml c
WHERE
/* --- Inclusion Criteria --- */
( /* --- Age >= 16 years and Age <= 60 years --- */
  CURDATE() - INTERVAL 16 YEAR >=
  DATE(EXTRACTVALUE(CCR_XML_String, '/ContinuityOfCareRecord/Actors/Actor[contains(ActorObjectID,.../Patient/ActorID) OR
  contains(ActorID,.../Patient/ActorID)]/Person/DateOfBirth/ExtractDateTime'))
  AND
  CURDATE() - INTERVAL 60 YEAR <=
  DATE(EXTRACTVALUE(CCR_XML_String, '/ContinuityOfCareRecord/Actors/Actor[contains(ActorObjectID,.../Patient/ActorID) OR
  contains(ActorID,.../Patient/ActorID)]/Person/DateOfBirth/ExtractDateTime'))
)
AND
/* --- Both male and female --- */
EXTRACTVALUE(CCR_XML_String, '/ContinuityOfCareRecord/Actors/Actor[contains(ActorObjectID,.../Patient/ActorID) OR
  contains(ActorID,.../Patient/ActorID)]/Person/DateOfBirth/ExtractDateTime'))

```

Local Gateways Query All Retrieve

Roles for taweela Physician Remote Query

Query < 21 >completed
(Remote) completed...

-1 Entered Query save Query Cancel

Returned Result

Clinic	Count	Latency Time	Error messages
	0		
	0		
	3		
	20606		
	1		
	490		
	20606		81.749 seconds
	20606		
	20606		
	11625		

Include Male and Female Patients with age >= 16 AND <= 60
that have ANY of the follow conditions:
-Active Diabetes Mellitus
-Active Diabetes Insipidus
AND ANY of the follow conditions:
-Active Anemia
==AND==
that have ANY of the follow lab tests:
-Serum LDL Cholesterol Measurement
-Serum cholesterol measurement >= 7.2
-Serum HDL Cholesterol Measurement
==AND==
that have ANY of the follow vital signs:
-O/E - blood pressure reading
-O/E - Diastolic BP reading >= 80.0 AND <= 100.0
-O/E - Systolic BP reading >= 120.0 AND <= 140.0
Exclude No Patients

Clinical Trial Data management Systems

Proprietary hell

Numerous partial solutions, open source forms are often quite basic in functionality

Little standardisation in spite of CDISC

Like buying a new CD player for every album

No 'one system' for every domain

The solution

Separate domain semantics from the system implementation.....

Requirements

A domain object model

A system that can author specific (xml) documents as an instance of the domain model for a specific study

A system to write the xml representation into a set of forms and a database

Define a new data element

The screenshot displays the ePCRN Study Designer v1.0 (Alpha) interface. On the left, the 'CRF Collection' pane shows a tree view of the study design. Under 'Form Collection 1', 'Page 1', and 'First Section', various data elements are listed, including 'Blood Pressure', which is highlighted with a red arrow. Below this is the 'Flow Rules Log' pane.

The main area is divided into two panes: 'CRF Description' and 'CRF UI Design'. The 'CRF Description' pane shows the details for the selected 'Blood Pressure' data element:

- Question Details:**
 - Name: Blood Pressure
 - Status: DRAFT NEW
 - Definition: Blood Pressure
 - Display Order: 7
 - Repetition: 0
- Metadata:**
 - ID: 645001412087757260
 - Version: 1.0
 - Date Created: 21-Apr-2008 12:13:24
 - Created By: [empty]
 - Context: [empty]
 - Origin: EPCRN

Buttons for 'Edit', 'Add/Find a CDE', 'Preview Form', and 'Design Form UI' are visible. Below this is the 'Information about This Question' section, which includes a 'DE Details' tab and a 'Question Flow Rule Details' tab. The 'DE Details' tab contains a list of buttons: 'Find CDE', 'Edit CDE', 'Delete CDE', and 'Create New CD'.

Use a data standards repository

The screenshot shows the ePCRN Study Designer v1.0 (Alpha) interface. The main window is titled "ePCRN Study Designer v1.0 (Alpha)" and has a menu bar with "File", "Edit", "Window", and "Help". Below the menu bar is a toolbar with icons for file operations. The main area is divided into several sections:

- CRF Designer UI**: Includes a search bar with "Enter Term:" and a "Search" button. Below it are radio buttons for "CDEs" (selected) and "Forms".
- Found Matching CDEs:** A list of search results. "Blood Pressure Systolic" is highlighted. A red arrow points to this entry.
- CDE Details:** A form showing details for "Blood Pressure Systolic".
 - Long Name: Blood Pressure Systolic
 - Status: RETIRED PHASED OUT
 - Definition: Systolic Blood Pressure
 - Public ID: 2004388
 - Version: 1.0
 - Context: DCP
 - Type: (empty)
- Information about this CDE:** A tabbed interface with "DE Details" selected. It contains two tables:
 - DE Details Table:**

Field Name	Value	Description
DE Long Name	Blood Pressure Systolic	
Preferred Defn	Systolic Blood Pressure	
Public Id	2004388	
Version	1.0	
 - Reference Documents Table:**

Document Name	Document Type	Document Text	Document Context
LONG_NAME	Preferred Question Text	Blood Pressure Systolic	DCP
- Filter:** Checkboxes for "Approved", "Released", "Retired", "Draft", and "Submitted".
- Messages:** A log showing system messages like "Concepts Loaded..." and "DSRSearch thread completed: Concepts loaded...15".

Create your own form design

The screenshot displays the ePCRN Study Designer v1.0 (Alpha) interface. The window title is "ePCRN Study Designer v1.0 (Alpha)". The menu bar includes "File", "Edit", "Window", and "Help". The toolbar contains icons for file operations and navigation. The main workspace is divided into two panes: "CRF Collection" on the left and "CRF Description" / "CRF UI Design" on the right.

CRF Collection Pane: Shows a tree view of the form structure. The root is "Form Collection 1", which contains "Baseline Form". Under "Baseline Form" is "Page 1", which contains "First Section". Under "First Section" are fields: "Name", "Gender", "Date of Birth", "HbA1c", "HDL", "LDL", "Diabetic?", "Blood Pressure", "Systemic", "diastolic", and "New Section 1.1.2". Under "New Section 1.1.2" are fields: "Q0", "Q1", "Q2", "Q3", "Q4", "Q5", "Q6", and "Q1.7". A red arrow points to the "Systemic" field under "Blood Pressure".

CRF UI Design Pane: Shows the form design for "Page 1". The form name is "Baseline Form". The form description is "Baseline Form (Status: DRAFT NEW) (Baseline Form)". The form contains the following fields:

- Name: Text input field
- Gender: Dropdown menu (Male)
- Date of Birth: Date picker (21/04/2008)
- HbA1c: Text input field
- HDL: Dropdown menu (50)
- LDL: Text input field
- Diabetic?: Dropdown menu (No)
- Blood Pressure: Grouped fields
 - Systolic Blood Pressure: Text input field
 - diastolic Blood Pressure: Dropdown menu (30)

PCROM and ISO11179

ISO 11179 is a six part international standard for meta data registries where each data element is

- Registered
- Uniquely identified
- Named
- Defined –for which a domain model is useful
- May be classified – for which a domain model is essential

Progress planned from 2010

ePCRN is an open source project

**€7M of funding from the European Commission
via the TRANSFoRm project
(www.transformproject.eu)**

Plans for widely usable tools for:

- Authoring searches using controlled vocabulary
- Working with metadata and forms
- Writing computable protocols

Where does the future lie?

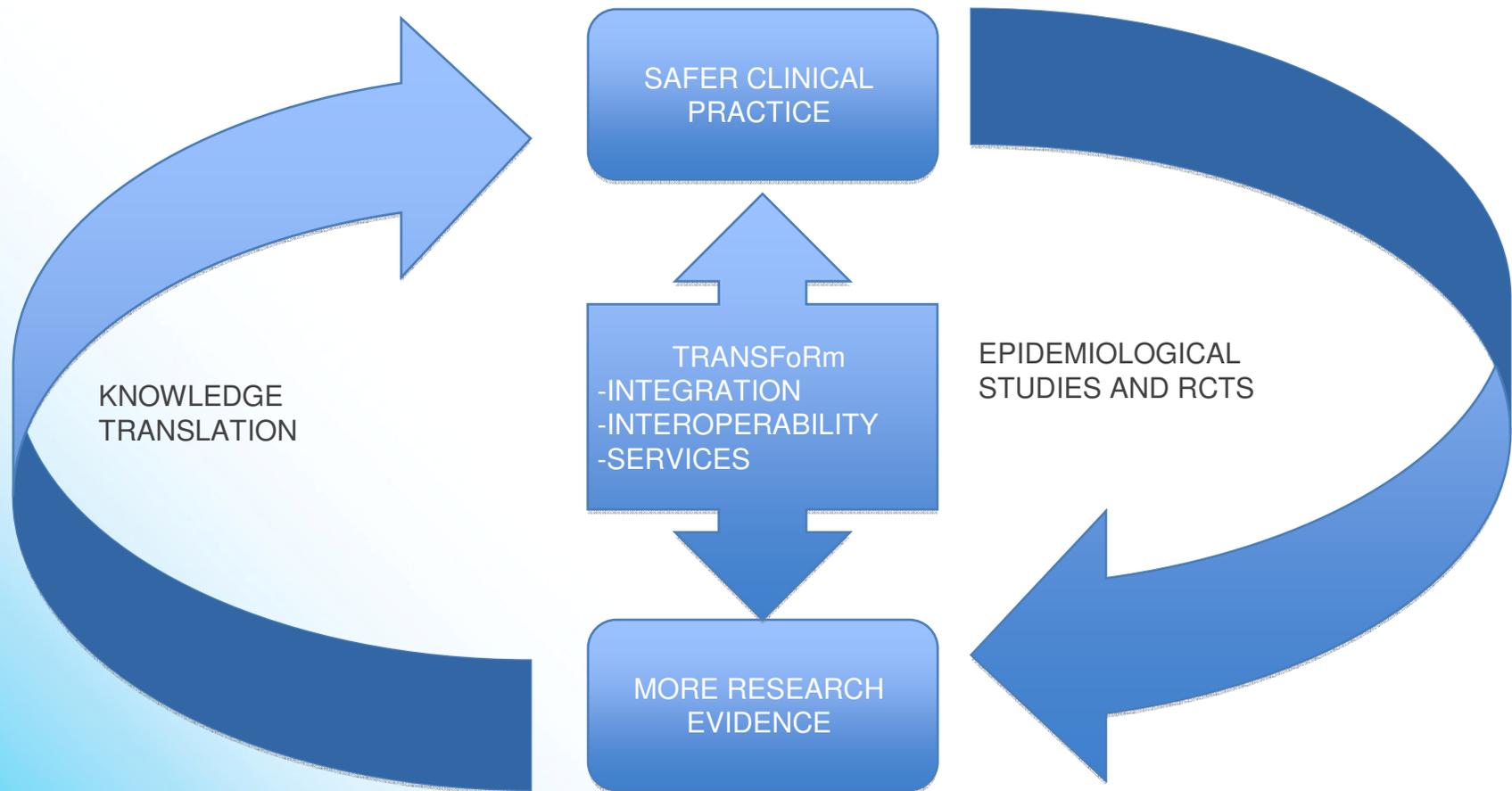
Open systems

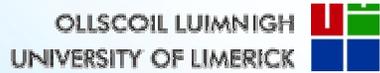
Standards

Standardized computable trial protocols

Ability deploy studies in a variety of different software systems

TRANSFoRm





Demonstration and dissemination

User requirements
Development and
Evaluation

WP 1
RESEARCH
USE CASES

WP 2
PATIENT
SAFETY USE
CASE

WP3

LEGAL
ETHICAL &
SECURITY
FRAMEWORK

WP4

DECISION
RULES AND
EVIDENCE

WP7

SYSTEMS AND
SERVICES FOR
DATA
INTEGRATION

WP5

USER AND
SOFTWARE
SERVICES

WP6

CONCEPTS
OF DATA
INTEGRATION

WP8

DEMONSTRATION

INDUSTRY
CONTRACT
RESEARCH
ORG
ACADEMIA

WP9

DISSEMINATION

WEBSITE
PUBLICATION
WORKSHOPS
PROTOTYPES
COLLABORATION
INFRASTRUCTURE
TECHNOLOGY
PLATFORMS

WP10 MANAGEMENT

ICT

System Integration

Using the models, based on use cases to inform the development of a system that has shared concepts:

- Provenance model
- Protect the data via provenance and common security model
- Consent model
- Middleware based on open source and standards (and building on ePCRN)

Workflow Integration

For research

- Seamless prompts, reminders and data collection

For clinical practice

- DSS have failed
- Strong emphasis on development from MDM experimental approach
 - Prompting v Alerting
 - Data entry interface and design

Knowledge translation

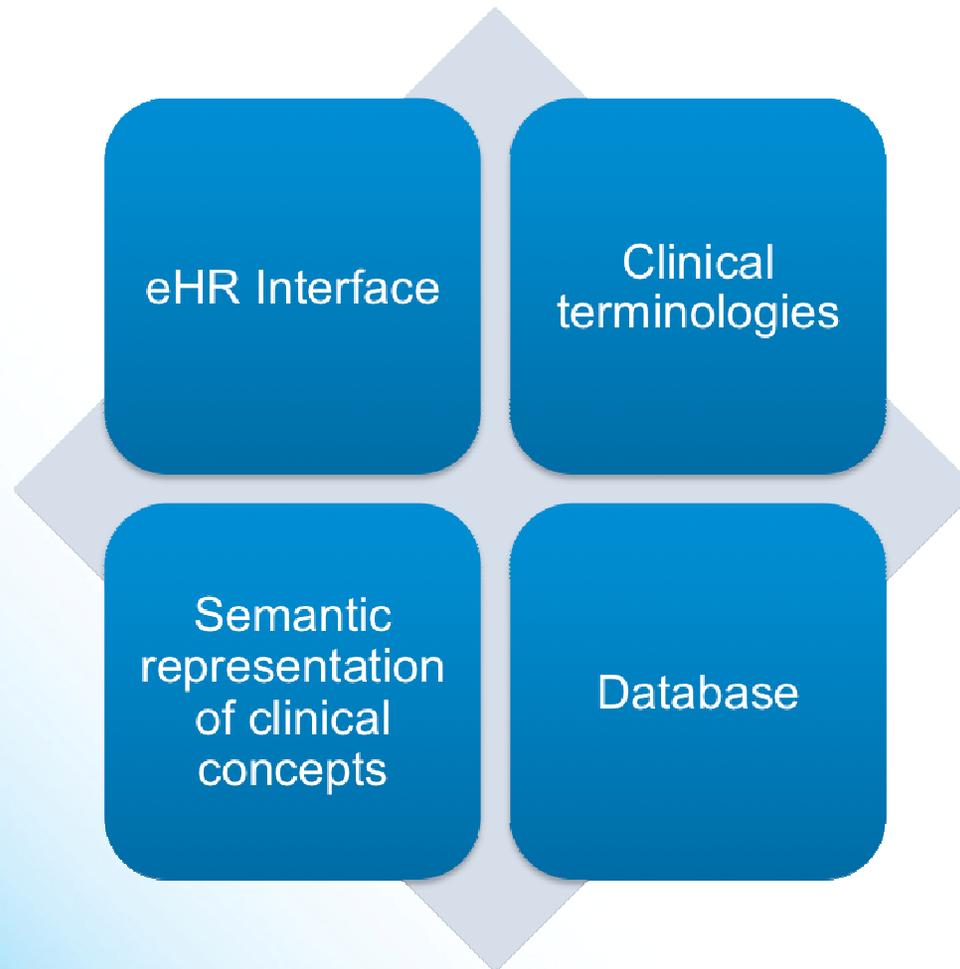
Understand how systems work with human judgment

Need to enable a much finer granularity of interaction between the eHR and research systems

Use the openEHR standard

- Observations, archetypes, templates
- User interface design

CEN 13606: independence of semantic representation.



Archetypes

A computable expression of a domain content model in the form of structured constraint statement based on a reference information model.

Often encapsulated together in Templates.

Sit between lower level knowledge resources and production systems

Independent of interface and system

Interoperability of research data

Model based

- CDISC ODM
- BRIDG
- PCROM

ISO/IEC 11179 Metadata Registry standard

Integrate Vocabulary service, Metadata registry and core interoperability model as a seamless service.

Current progress

Use case development

Experimental background to DSS

Survey of network capacity and IT resource

Model of data comparability

Legal and ethical survey

Provenance framework

Security Framework

QA Framework

Potential areas for co-operation

Standards

- Clinical data representation/ archetypes
- Standard computable representation of clinical research / trial data

Common Implementation

- eHR interface
- Middleware / system integration

Common Business Models

- Open but exploitable
- epSOS and EHR4CR projects

Overcoming the barriers

Funding for research

- FP7, ARRA, CTSA inhibit 'ePCRN style' collaborative research

Funding for networking and collaboration

- Goes beyond individual projects and needs LONG TERM investment

Involvement of standards bodies

- Cannot be a short term aim

Agile and open

- Should encompass all those working in the area