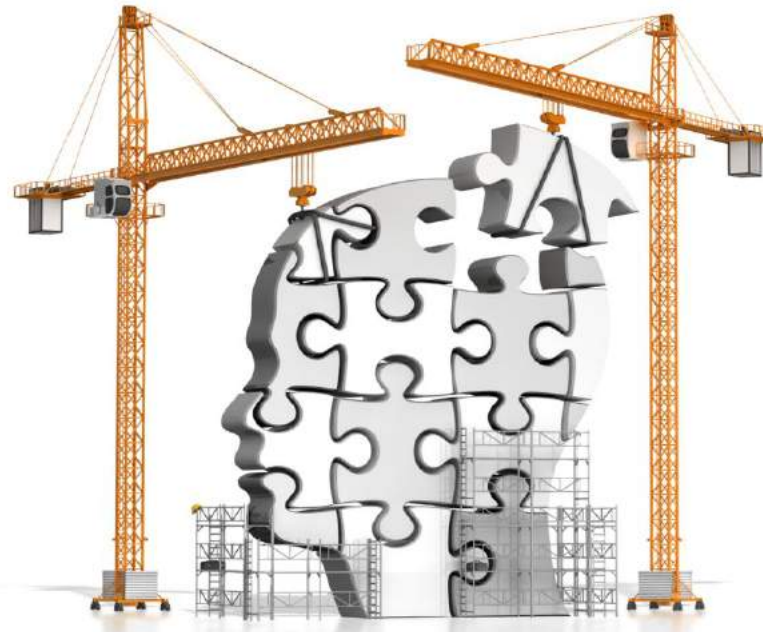# Building infrastructure for digital health using open innovation 2.0 in co-design of a meta data registry framework to support integrated care

Dr. Pamela Hussey, Dr. Kris McGlinn

Overview

- Introduction to the Centre for eIntegrated Care (CeIC) and Integrated Architecture (IA) and its goals

- What is the Health and Social Care National Data Dictionary (HSCNDD) meta data registry and why is it needed?

- How can emerging semantic web technologies support the development of a national data dictionary

- Scoping of standards in the health domain in Europe and Internationally

- An ontological approach to an Irish National Data Dictionary based on an analysis of Irish data dictionaries.
  - With a focus on representing demographic data

About the CeIC

- Strong background in health informatics standards through NSAI HISC, ISO TC215 and CEN TC251

- Accredited R&D Centre for International Classification of Nursing Practice (ICNP©) since 2016

- Core research interests Health Informatics, Concepts and Terminology, Interoperability , Electronic Health Record

Trinity College Dublin

eHealth Ireland

CeIC

What is Interoperability?

Is the ability of two or more systems or components to exchange information and to use the information that has been exchanged (IEEE1990)

It includes:

1. The exchange of information (Technical)

2. The ability of the recipient to use that information (Semantic)

3. The business methods to enable interoperability (Process interoperability)

4. Clinical practice and utility (Clinical interoperability)

Benson and Grieve , 2016

About the Information Architecture (IA)

- The IA is part of the Enterprise Architecture Authority in eHealth Ireland
  - see: http://www.ehealthireland.ie/Our-Team/Enterprise-Architecture/Information-Architecture-/

- Its primary concern is to develop metadata (information about data) to improve interoperability between IT systems

- IA publishes its metadata through catalogues
  - see http://hive-worx.com/indd/

- **CeIC is developing a metadata catalogue** for the National Health and Social Care Data Dictionary (HSCDD)
  - HSCDD is intended to function as a single source of truth (SSOT)

National Health and Social Care Data Dictionary (HSCDD) Overview

- Innovation 2.0 Quadruple helix ecosytem

Financing
Policy formulation
Innovation
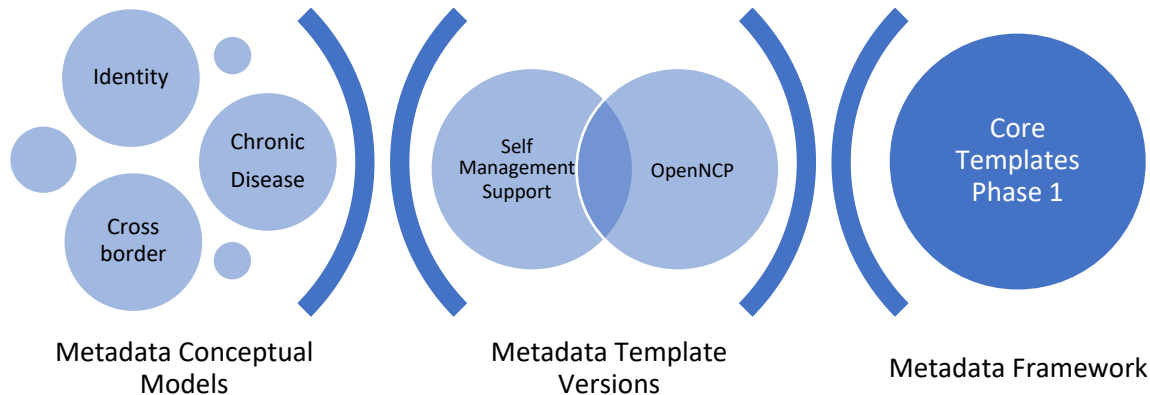Support
Advisory services
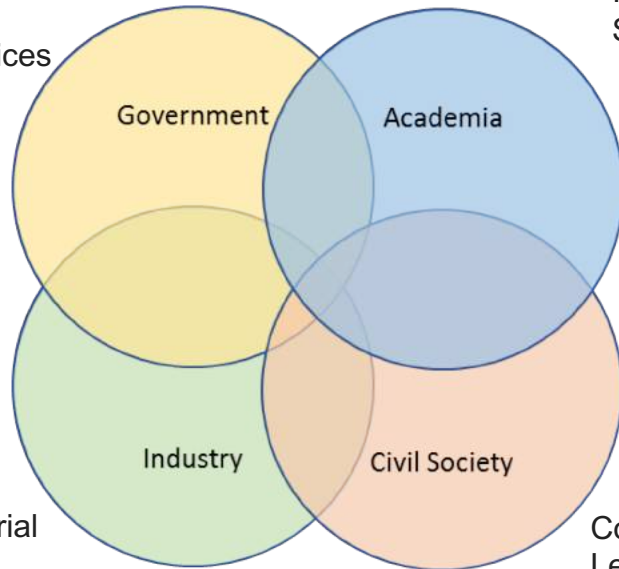
R&D
Education
Incubators
Spin-offs

Government

Academia

Industry

Civil Society

Entrepreneurial
Venturing
Product and
service
development

Collaboration platform
Legitimating link
Competence & process
development
Carrying individual and
structural perspectives

National Data
Dictionary
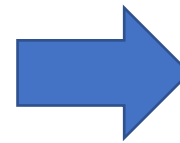
"centralized repository of information
about data such as meaning,
relationships to other data, origin,
usage, and format".

Meta Data Registry Framework Big Picture

Identity

Chronic Disease

Cross border

Self Management Support

OpenNCP

Core Templates Phase 1

Metadata Conceptual Models

Metadata Template Versions

Metadata Framework

One initiative we are exploring is to use Semantic Web technologies to manage interoperability!

Review of Evidence and Discussion Papers
http://www.ehealthireland.ie/Our-Team/Enterprise-Architecture/Information-Architecture-/
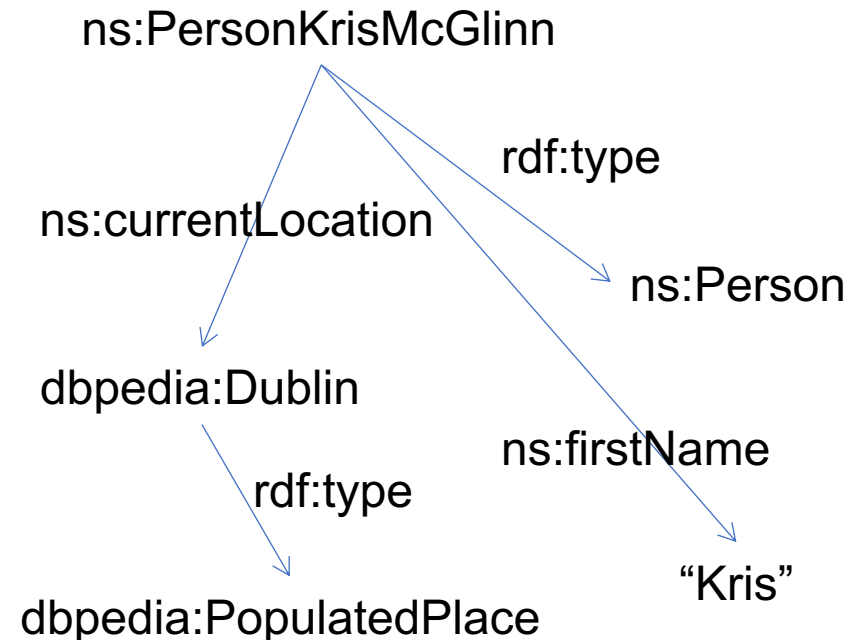
A brief introduction to the Semantic Web

- What is it?
    - Essentially it is the Web of Data.

Web of Documents   VS   Web of Data

ns:PersonKrisMcGlinn

rdf:type

ns:currentLocation

ns:Person

dbpedia:Dublin

ns:firstName

rdf:type

dbpedia:PopulatedPlace

"Kris"

A brief introduction to the Semantic Web
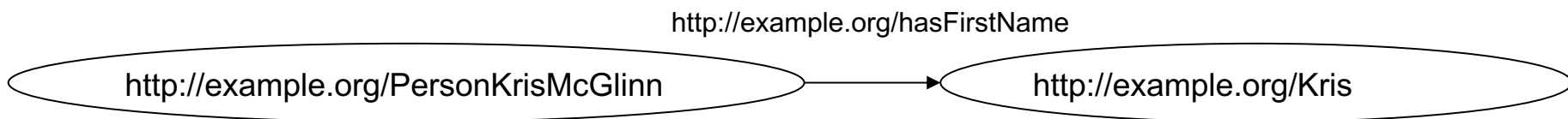
- What is it?
  - Essentially it is the Web of Data.

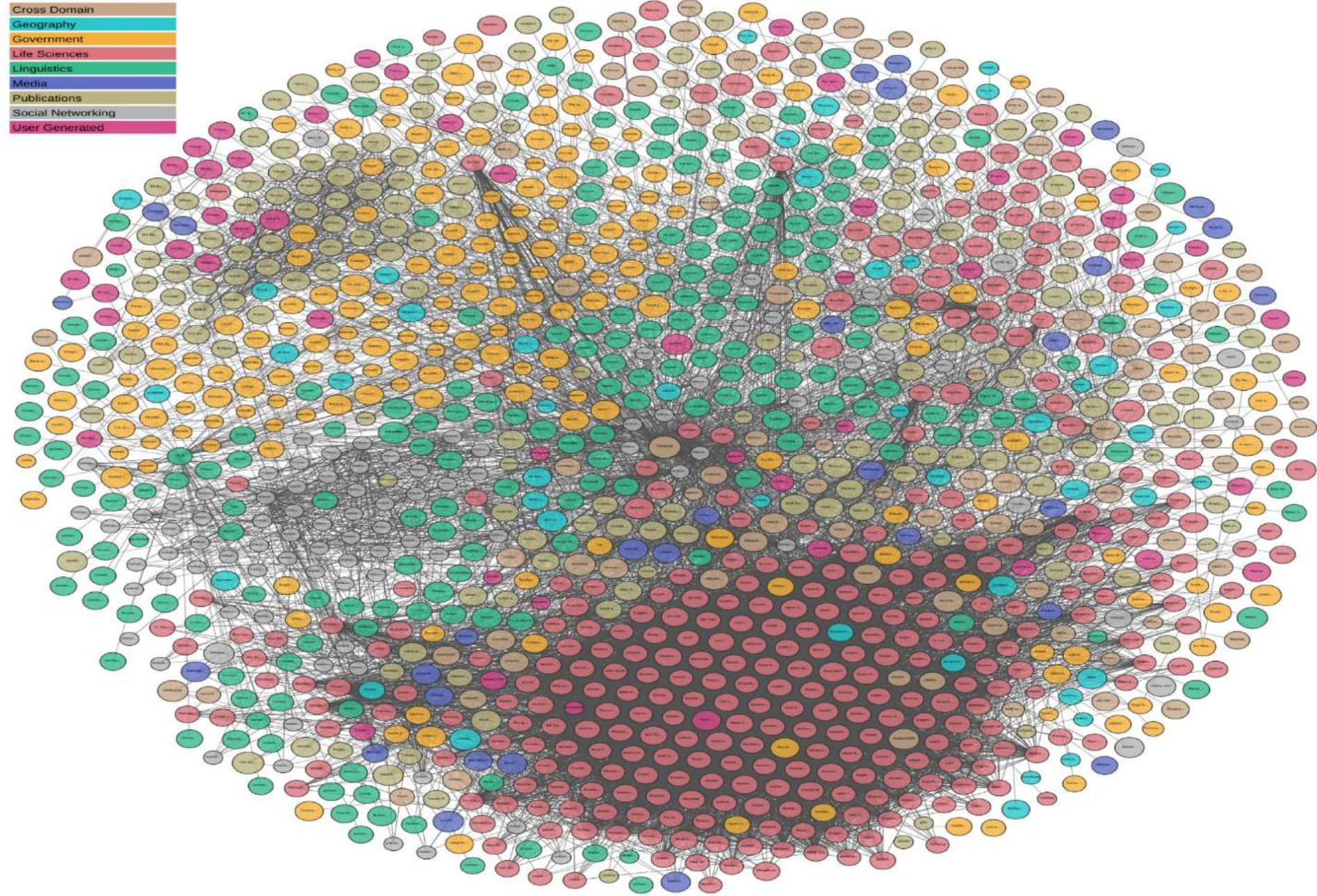 "Semantic Web Technologies" is a collection of standard technologies to realize a Web of Data

- Of course, the devil is in the details
  - a common model has to be provided for machines to describe, query, etc, the data and their connections
  - the "classification" of the terms can become <u>very</u> complex for specific knowledge areas: this is where ontologies, thesauri, etc, enter the game…

Core Idea: Resource Description Framework

- A scheme for defining information on the web.
  - It provides the technology for expressing the meaning of terms and concepts in a form that computers can readily process.

- RDF encodes this information in sets of triples.
  - The triple is information on the web about related things.

- Each triple is a combination of Subject, Predicate and Object, similar to a elementary sentence
  - each identified by a URI
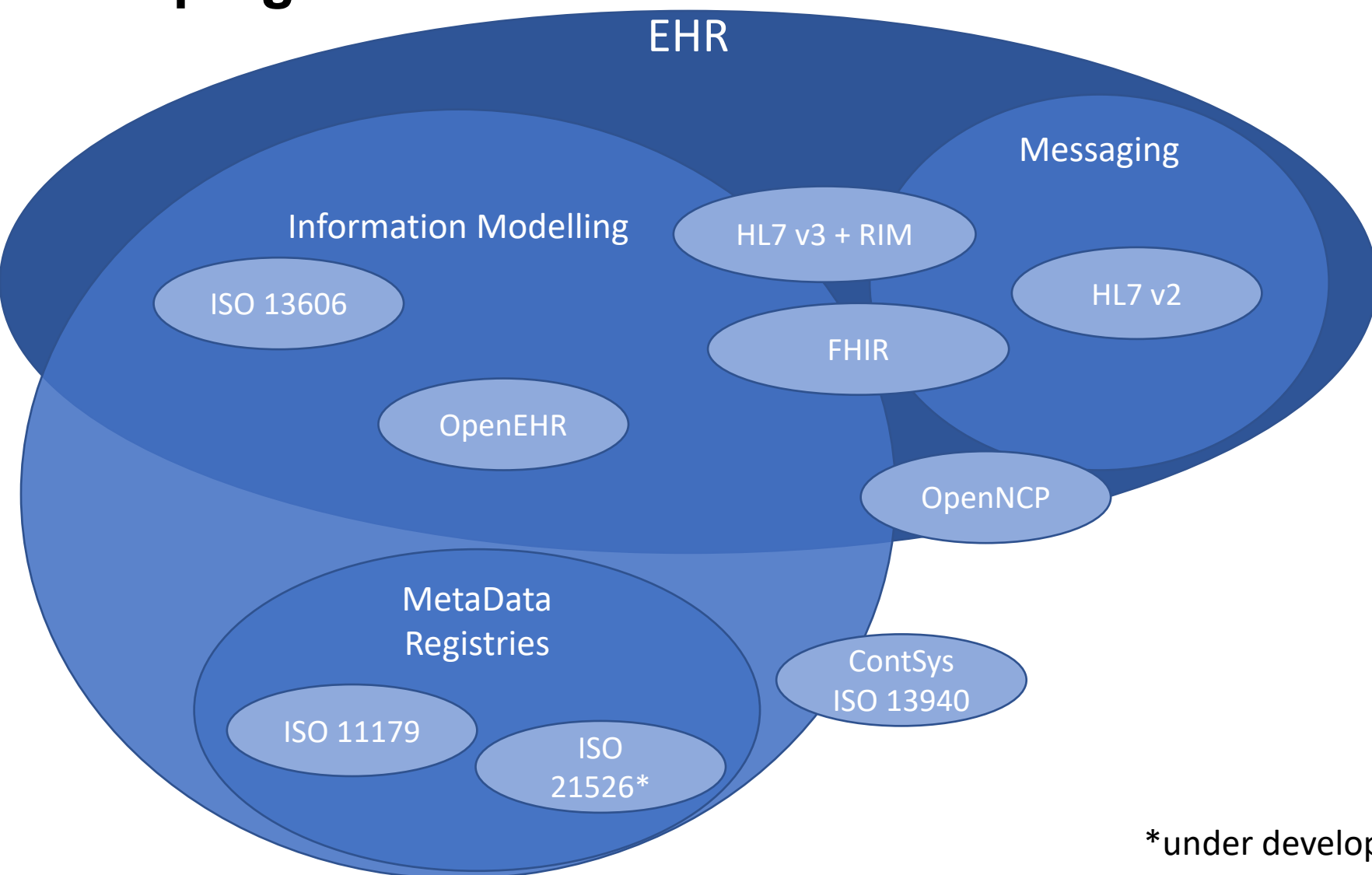 - enable anyone to define a new concept just by defining a URI for it somewhere on the web.

http://example.org/hasFirstName

(http://example.org/PersonKrisMcGlinn) ⟶ (http://example.org/Kris)

The Linked Open Data Cloud - https://lod-cloud.net/

Trinity College Dublin

eHealth Ireland

Scoping Standards in Health Domain

EHR

Messaging

Information Modelling

HL7 v3 + RIM

HL7 v2

ISO 13606

FHIR

OpenEHR

OpenNCP

MetaData Registries

ContSys ISO 13940

ISO 11179

ISO 21526*

*under development

Trinity College Dublin

eHealth Ireland

CelC

Metadata Registries in Healthcare

- ISO/IEC 11179 Metadata Registry (MDRMetamodel)

- ISO/DTS 21526 MetaRep
  - extends ISO/IEC 11179, as well as ISO/JTC1/SC32/WG2
  - core of MetaRep are the common facilities which in total define 24 classes
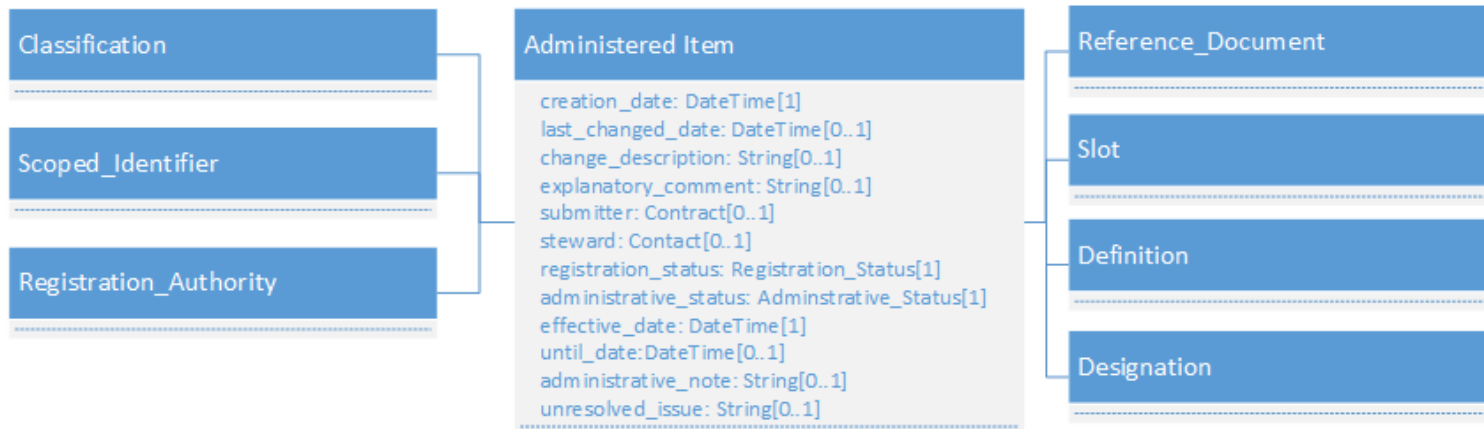
Analysis of Irish Healthcare Data Dictionaries

- The Health Information and Quality Authority (HIQA) generates a catalogue of national health and social care data collections

- Provides a list of 75 collections with data in terms of
    - title, managing organisation, description/summary, data providers, available data dictionaries, **data content** (i.e. a breakdown of the type of data collected) etc.

- HIQA has also generated the National standard demographic dataset and guidance for use in health and social care settings in Ireland
    - provides guidelines on a set of concepts and properties for describing demographic data, such as related to name, date of birth, contact details, address, etc.
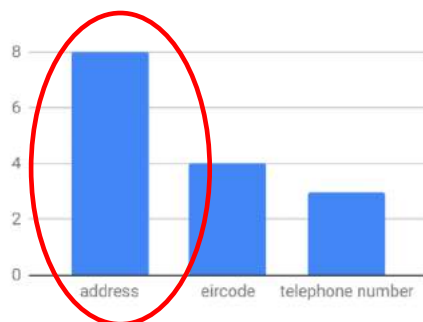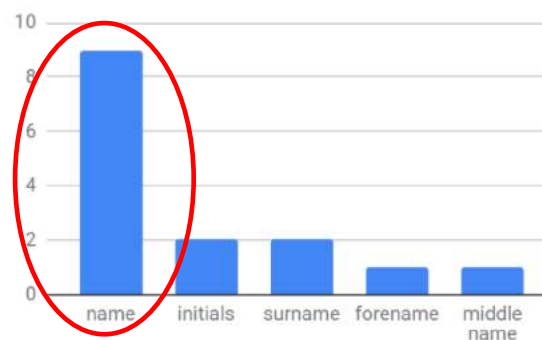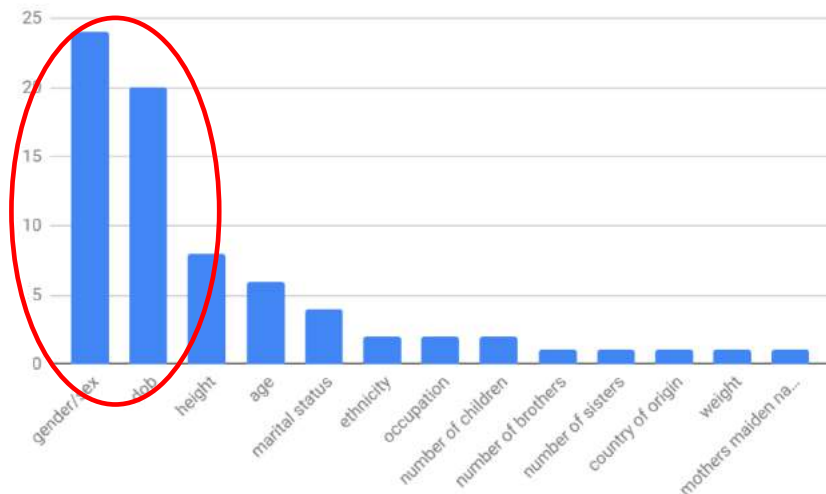
Analysis of Irish Healthcare Data Dictionaries

- Methodology:
    - Step 1: the data dictionaries given in the data dictionary field were analysed.
    - Step 2: different named concepts were extracted from the data content field.
    - Step 3: harmonisation, to identify a set of classifications for the different concepts identified.

- Results Step 1 Data Dictionary Analysis:
    - 39 had "no", "not available", or "not available online".
    - remaining 36, 15 provided links (such as www.noca.ie) with no explicit way to access the data dictionary or required a password, 3 had broken links, 5 mentioned resources which could not be located (e.g. Under revision as part of HRB LINK project) and so these were discounted.
    - The remaining 13 data dictionaries were mostly pdf documents, such as the ACR, CFR, PCR and PTR standards, as well as EUROCAT, and heartwatch. The Irish Mental Health Care provides an excel file.

Results Step 2 Concept Extraction from Data Field

- 75 collections "data content" field was analysed

- typical examples of this type of data (without a corresponding data dictionary)
  - "Name, address, date of birth, gender, District Electoral Division (DED), HSE area, Local Health Office (LHO) area, task force area, date commenced on methadone, type of methadone treatment, prescribing doctor, dispensing clinic, date and reason for discontinuation of methadone, client photograph and client signature."

- a matrix of collections against listed data concepts was created
  - a tick was given for a data concept if it is present in a collection

- over 50 potential classes have been identified within the Irish health domain, ranging from (not an exhaustive list):
  - person, name, contact details, patient, patient infant, patient pregnant, disabled person, address, location, medical/clinical information/assessment, treatment, therapy, prescription, observations, test, results, diagnosis, event, incident, injury, death, paediatric mortality, service, procedure, operation, product, device, vehicle, vaccination/immunisations, disease, infection, staff, practitioner, admission, child admission unit, legal status, approved centre, etc.
  - Typical data concepts are those related to demographic data on patients (i.e. age, gender etc.).

Results Step 3 Harmonisation (Demographics)

A Data Dictionary Ontology for the Irish Health Domain (Demographics)

*https://www.scss.tcd.ie/~mcglink/ontologies/metarep.html
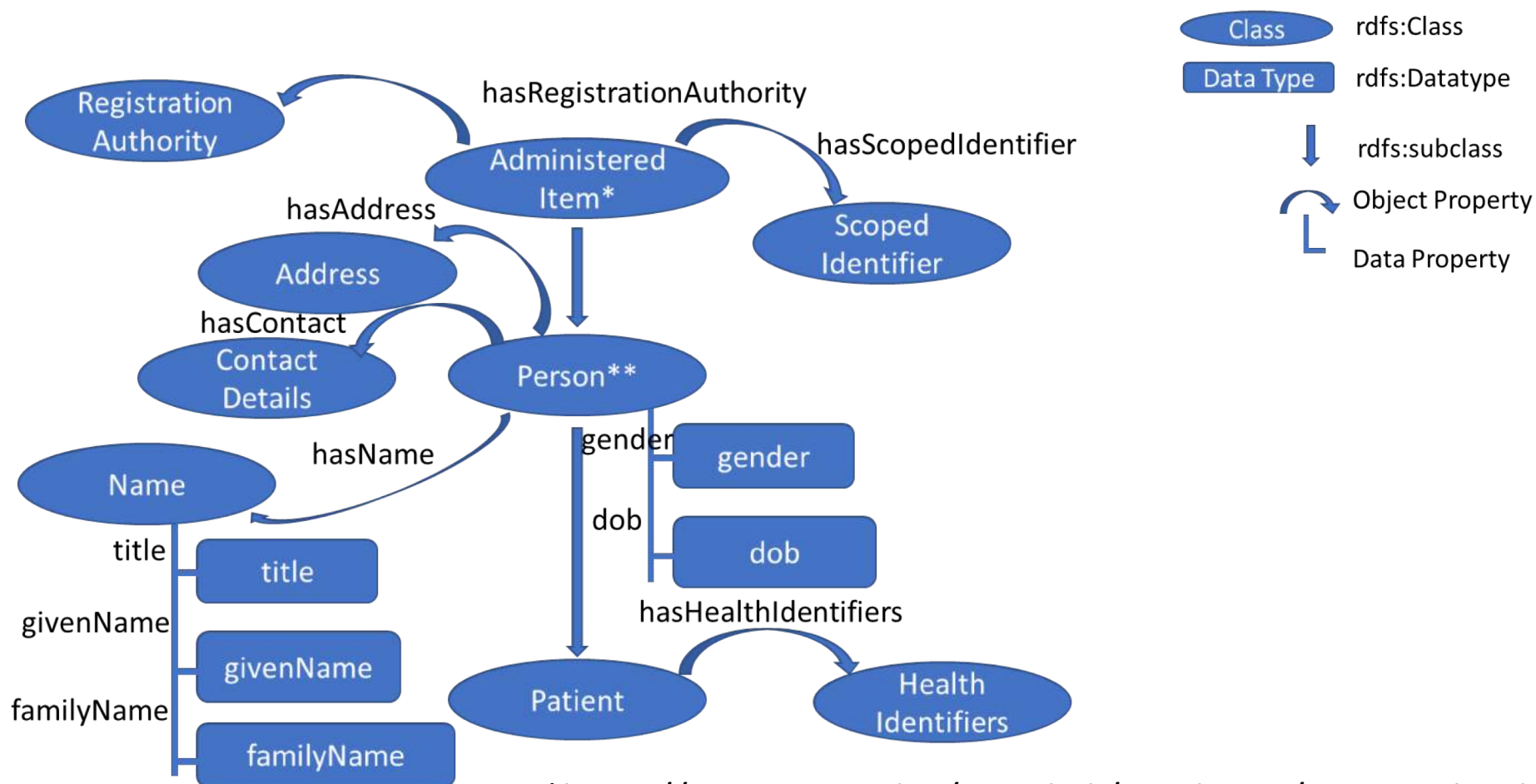
**http://trajano.us.es/~isabel/EHR/Demographic_RM.owl

Next Steps

- Continue the iterative development of the ontology

- Explore the integration of PROMs and PREMs

- Begin exploring data uplift into semantic representation
  - E.g. conversion of tabular data into RDF

Question?

- Thank you