CO-ODE
Introduction & Starting point for questions

Alan Rector
Bio and Health Informatics Forum/
Medical Informatics Group

Department of Computer Science
University of Manchester

rector@cs.man.ac.uk
www.cs.man.ac.uk/mig img.man.ac.uk
www.clinical-escience.org
mygrid.man.ac.uk

OpenGALEN

BHIF

CLEF

MRC
Medical Research Council

Projects - Officially

- **CO-ODE – Cooperative Open Ontology Development Environment**
  - **Protégé + OilEd/OWL + Features from PCPACK & GALEN**
    - **Funded as part of the UK E-Science Infrastructure by JISC the same people as fund JANET and joint libraries**
      - **2 years x 2 people + subcontracts to Stanford and Southampton**

- **Part of a larger team on the semantic web, ontologies, and bioinformatics**

- **HyOntUse – Hybrid Ontology Develement for Usabiity**
  - **Debugging aids & views for ontologies**
    - **Funded as a companion project by EPSRC (NSF equiv)**

OpenGALEN BHIF CLEF MRC Medical Research Council

A Special Opportunity

- **Parallel funding on both sides of the Atlantic**
  - **UK: JISC + EPSRC**
  - **US: National Cancer Center Center for Bioinformatics (NCICB)**
    - **(not to be confused with the NCBI)**

- **The two most widely used tools + special expertise on interface, elicitation, & transformations**
  - **Both with user bases in the thousands**

- **The two key paradigms for knowledge representation**
  - **Description Logics/OWL and Fames**

To build something special

*Open*GALEN   BHIF   **CO-ODE/HyOntUse**   CLEF   MRC

Who in Manchester

- **Health and Bioinformatics Group/Forum**
  - **Healthcare records, terminologies, and decision support –**
    - **PEN&PAD, GALEN, CLEF - Alan Rector, Jeremy Rogers..**
  - **Bioinfoinformatics & BioOntologies**
    - **Gene Ontology Next Generation/Gene Ontology Annotation Tool (GONG & GOAT) – Robert Stevens, Chris Wroe, Mike Bada**
    - **Gene Ontology Annotation &**

- **Information Management Group**
  - **E-Science & Bioinformatics**
    - **^{my}Grid – Carole Goble, Norman Paton, & a cast of thousands**
      - **Many other onotology users: Astronomy, Aircraft industry, Art history, …**
  - **Semantic Web**
    - **WonderWeb/WebOnt, – Ian Horrocks,**
    - **OWL Infrastructure – Sean Bechhofer**

*Open*GALEN

BHIF

CLEF

MRC
Medical Research Council

Some of what we do

- **Build applications using ontologies**

- **Develop metadata for web services**

- **Build ontologies & terminologies**

- **Build ontology development environments**

- **Build "ontology servers" and developing "ontology services"**

*Open*GA LEN

BHIF

CO-ODE/HyOntUse

CLEF

MRC
Medical Research Council

We Need

- **A comprehensive ontology development environment**
  - **That will scale to build big ontologies**
  - **That's part of a community so others build tools**
  - **That can be used to build knowledge on ontologies**

- **An application development platform**

- **User friendly tools with sound semantics**
  - **Used by biologists, doctors, engineers, astronomers, NOT just by ontologists**
    - **But quick and easy for professional ontologists as well.**
  - **Description logics as the "assembly language" of ontologies**

We Need(II)

- **Tools to help people build ontologies rationally**
  - **Tools to help people get started in new domains**
  - **A version that encourages"normalised ontologies"**
    - **A version that encourages understanding metadata**
  - **Support for development by transformation**
    - **Formalise what exists - GO, FMA, Mouse Anat, MGED, UMLS…**
      - **Reduce hand crafting**
    - **Extensions of existing ontologies**

- **Tools for application building**
  - **Applications are the goal; ontologies the means; and DLs the means to ontologies**
    - **Hybrid systems**

- *A de facto standard*
  - *The two most widely used tools in ontology development*
    - *Each has a user base of several thousand*
      - *Although there are lots of others around*

OpenGALEN   BHIF   **CO-ODE/HyOntUse**   **MRC** Medical Research Council

Customers (Academic UK)

- ## Local
  - **Chris Wroe and Robert Stevens - The Gene Ontology**
  - **CLEF – knowledge resources and electronic health records**
  - **Other bioinformatics**

- ## Other Bioinformatics
  - **The EBI -MIAME/MGED for Proteome and ??? for metabolome …**
    - **Alternative Editors: DAG-EDIT, ECLIPSE, GOET**
  - **Radiologists wanting to describe both method and images – MIAS-GRID**
  - **National Cancer Research Institute (NCRI), National Cancer Translational Network (NTRAC), National Caner Tissue Resource (NCTR) Sylvia Nagle**
  - **Combined mammalian anatomy project & SONNET– Human + Mouse Duncan Davidson, David Shotton, Cornelius Rosse, …**

- ## The E-Science Grid Commnity
  - **^{my}Grid, Geodise, AstroGrid, MIAS-Grid, AKT, …**

- ## The knowledge management / data curation / information environment community
  - **Library / digital library community**

Customers (Commercial)

- **Big Pharma**
  - **GSK - Robin MacIntyre / Howard Bilofsky**
    - **Ontologies for metadata**
    - **Forms management for Clinical Trials**
    - **Pheonotype to genotype links**
    - **Invitation / Command performance to visit Philadelphia early in project**
    - **(The have taken myGrid alpha for testing!)**
  - **Astra Zeneca – Anne Westcott**
    - **Trialing information**
    - **Information extraction**
    - **Wants real collaboration**

- **Knowledge Management**
  - **CSW – manage parts of BMJ and BMJ Evidence and much of National Electronic Library for Medicine**

- **Health IT**
  - **Skeptical interest…**

- **Network Inference ? Epistemics ?**

OpenGALEN

Customers (Paying) -

- **JISC – a tool for the E-Science Community**
  - **Other projects**
  - **Special groups**
    - **Curation centre and UKOLN – librarians and curation of primary data**
    - **Text extraction centre**

- **NCICB – Ken Buetow, Frank Hartel**
  - **Funding Stanford**
  - **Pushing for influence & role with NCRI in UK**
  - **Open source tool for their ontology development**
    - **A big win**

OpenGALEN

CO-ODE/HyOntUse

Complementary Functionality
Why we need both Frames and OWL

- **OWL/DLs – an 'ontological core' for Classes/Types (The T-Box)**
  - **Essential knowledge – the "ontology proper"at the kernel**
  - **Managing complex Subsumption lattices (DAGs)**
    - **Keeping track of multiple parents by hand doesn't work**
  - **Reasoning and expressiveness**
    - **Saying more of what you want to**
      - **negation, disjunction, some/all/cardinality**
    - **Open World – negation as unsatisfiability**

- **OilEd Infrastructure – a rapidly evolving architecture**
  - **Change management and evolution**
  - **Provenance**
  - **Practical demonstration of how to do it**

*Open*GALEN   BHIF   **CO-ODE/HyOntUse**   CLEF   MRC Medical Research Council

Complementary Functionality
Why we need both Frames and OWL

- **Protégé /Frames– 'an ontology based KB' –** (Generic Knowledge, Supplementary Reasoning, Metadata, & Instances)
  - **Fixed asserted ontology – (we want to change that)**
  - **"Contingent Knowledge" The knowledge base around the ontology kernel**
    - **Defaults and Exceptions – classic frame paradigm**
  - **Metadata**
    - **Flexible, user defined, fine grained (on each statement & concept)**
  - **Supplementary reasoning and transformations**
    - **Things outside the DL fragment (or current DL reasoners)**
    - **Things more efficiently done in other ways – other reasoners**
  - **Instances**
    - **Most instance uses cases do not need a full open-world A-Box**
  - **Plug & Play architecture**
    - **Plugins – PROMPT, constraints, UMLS, JESS, Visualisation, … …**
    - **Applications!**

OpenGALEN

What's Missing

- **Usability and requirements studies**
  - *Interfaces the result of history*

- **The knowledge elicitation phase**
  - *Both assume preliminary analysis*

- **Easy starting points**
  - *Both are perceived as too complicated*

- **Updating to the web age**

OpenGALEN BHIF **CO-ODE/HyOntUse** CLEF MRC Medical Research Council

Hi Level Requirements

- **User Interface(s)**
  - **Intuitive to users from both frames and DLs**
  - **Help users formulate models**

- **Representation**
  - **Flexible metadata using Protégé**
    - **Each 'statement' may need its own provenance**
  - **Generic information**
    - **defaults and exceptions in frame paradigm**
    - **Extensions to DL paradigm using other reasoning**
  - **Instances**

- **Architecture**
  - **Use/adapt Protégé plugins – PROMPT, PAL, JESS, Visualisation,…**
  - **'Simple'/familiar Protégé view of DL based ontology/KB**

- **Transformations, "Intermediate Representaiton" & 'Meta authoring'**
  - **To simplify interface to user**

OpenGALEN

BHIF

CO-ODE/HyOntUse

14

CLEF

MRC
Medical Research Council

The Story So Far

- **One exchange of visits with Stanford**
  - **Holger Knublauch appointed at Stanford – an prototype "OWL Tab" nearly developed**
    - **Not yet attached to classifier**
  - **"New OilEd Infrastructure" well on its way in Manchester – High level**
  - **Discussions with Jena under way**
  - **The first layer of issues raised – and partly solved**
    - **Linking frames and DLs is NOT simple**
  - **RAs in Manchester appointed, to start in September.**

OpenGALEN BHIF **CO-ODE/HyOntUse** CLEF MRC

Commitments from Proposal

- **Requirements & Formative Evaluation**
  - **User requirements workshops or equivalent(2)**
    - **Preferably piggy-backed on other events**
      - **It is late, but do we want to try for anything at the All Hands meeting?**
    - **Discussions with specific projects**
      - **GO, MGED, Astrogrid, others?**
    - **Use of NESC?**
  - **User groups and lists**
    - **Protégé, OilEd, Semantic Web**
  - **Commercial consultation**
    - **GSK - when?**
    - **Others – who? how?**

- **Formal documentation of Formative Evaluation/ Requirements process**

*Open*GALEN   CO-ODE/HyOntUse   MRC

Practicalities

- **Collaboration mechanisms**
  - **Tasks & Dependencies**
  - **Optimal timing given 'events'**
  - **Points of leverage**
    - **How to make the most of limited resources**
  - **Coordination mechanisms**
    - **Phone conferences, video conference, Access Grid**
      - **(Stanford has no Access Grid)**

- **Public Face of Project**

- **What's missing?**

OpenGALEN   **CO-ODE/HyOntUse**