

Meeting report series

Report of the 4th DSC WG on Ontologies and Disease Prioritization Teleconference

11 March 2015

Participants

Dr Peter Robinson, Berlin, Germany (Chair)

Dr Kym Boycott, Ottawa, Canada

Prof Melissa Haendel, Portland, USA

Prof Ada Hamosh, Baltimore, USA

Ms Henrietta Hyatt-Knorr, Bethesda, USA

Ms Janine Lewis, Bethesda, USA

Ms Suzanna Lewis, Berkeley, USA

Dr Ana Rath, Paris, France

Dr Yaffa Rubinstein, Bethesda, USA

Dr Ségolène Aymé, Scientific Secretariat

Dr Lilian Lau, Scientific Secretariat

Ms Antonia Mills, Scientific Secretariat

Apologies

Prof Michael Bamshad, Seattle, USA

Dr Michael Brudno, Toronto, Canada

Dr Helen Firth, Cambridge, UK

Dr Chris Mungall, Berkeley, USA

Dr Helen Parkinson, Cambridge, UK

Prof Pak-Chung Sham, Hong Kong, China

Dr Damian Smedley, Cambridge, UK

Dr María Sobrido, Santiago de Compostela, Spain

Prof María Taboada, Santiago de Compostela, Spain

Ms Sharon Terry, Washington DC, USA

Agenda

- ▶ Update on the reorganisation of IRDiRC Working Groups
- ▶ Review of previous work/product of the Ontologies WG
- ▶ “IRDiRC Recommended” for WG products
- ▶ Discussion of need for an Ontologies Task Force in 2016

REPORT

Update on the reorganisation of IRDiRC Working Groups

As IRDiRC Working Groups (WGs) have fulfilled their mandates and provided their recommendations to the Executive Committee (Exec Comm) for action, this has led to a switch from the concept of WG to Task Forces (TFs) to implement some of these ideas.

A few ideas have been selected for action in 2015: Matchmaker Exchange, Patient Relevant/Reported Outcome Measures, Computable Consent Forms and Small Population Clinical Trials. More TFs will be set up for actions to implement in 2016 and beyond. IRDiRC is keen to invest on actions and welcome further proposals, subject to the approval of the Exec Comm.

Review of previous work/product of the Ontologies WG

ICHPT's Phenotype Ontologies

The work on phenotype ontologies by the ICHPT is close to completion, with a few more terms which need to be checked. Once completed, these terms will be published on the IRDiRC website, cross-referenced to other nomenclatures. This will be the first product of the WG. Additionally, there is a plan to document the definition of every term. Follow-up work to maintain the ICHPT minimum set of common terms may be scheduled to a TF to ensure necessary support.

Global Alliance Phenotype Ontologies WG

The Phenotype Ontologies Task Group (PO-TG) at the Global Alliance for Genomics and Health (GA4GH) also involves a few members of the IRDiRC Ontologies WG. One of the primary goals of PO-TG is to identify the ontological structures required for rare diseases as well as common-complex diseases, including cancer. PO-TG is currently working with the Genotype2Phenotype Association TG (G2P-TG) to explore the technologies and standards development involved in content sharing and how to represent phenotype data.

GA4GH works under similar principle as IRDiRC but goes beyond rare diseases to include complex diseases. The IRDiRC Ontologies WG and PO-TG should be merged or aligned in some way to avoid reinventing the wheel and repeating certain work. There is a lot of commonality and it would be ideal to model work in a fully interoperable system and one that could support data exchange.

NIH BD2K Workshop on Community-based Data and Metadata Standards Development

NIH recently organised a workshop which involved clinical standards and basic research community, many involved in phenotyping and other biomedical standards development. The focus of the workshop was on technical, social and financial aspects in developing standards and interoperability across the

translational spectrum. A report is currently in preparation, and it would be worthwhile to follow up on their activities given the overlap in scope.

“IRDiRC Recommended” for WG products

IRDiRC has just launched a new process to label databases, tools and recommendations to promote their use in rare diseases research. The purpose of “IRDiRC Recommended” is to help the rare diseases community to work optimally by bringing to their attention a set of useful aids for their research. Further information is available on the IRDiRC website (http://www.irdirc.org/?page_id=4474).

The ICHPT’s list of phenotype ontologies is an obvious candidate for the “IRDiRC Recommended” label, and the project leaders should apply for this recognition when the list has been made available to the community. Other tools including HPO and ORDO are encouraged to apply for the label. The label is not restricted to only ontology-based tools; any useful tool, database or registry is eligible to apply, e.g. OMIM, Phenotips, GeneMatcher.

Discussion of need for an Ontologies Task Force in 2016

TF on “Ontology Interoperability and Translation”

A proposal for a TF of “Ontology Interoperability and Translation” will be made to the Exec Comm.

The HPO has been receiving requests for translations of the phenotype ontologies into other languages (e.g. Spanish, French, Chinese and Japanese) but additional funding is required to support this work. The requests for translation could stem from interests of doctors or clinicians, not just researchers, to use bioinformatic solutions to improve clinical cares.

Orphanet has indicated its interest in playing a major role in the translation work. A plain language translation should also be considered to facilitate information transfer to individuals (e.g. patients) who may not understand the complex terms. Similar translation-related work has been planned at the NIH but lacks funding for execution.

One way forward would be to have a workshop that focuses on the process to facilitate ontology interoperability and translation. Its product will be a white paper or position statement that defines the needs, identifies the appropriate funding for such a project, and maintains its sustainability.

A draft paper previously circulated, and currently on hold, discusses the problem of interoperability and dissemination outside the context of the usual databases and publication venues. Tools should be created to interoperate not only limited to ontologies but also between registries, wikidata and other databases. There is a demand from undiagnosed patients to identify another patient and be matched.

The goals for such a workshop need to be better thought out. It is a unique vision and opportunity to assess the requirements for interoperability, minimal shareable phenotype profiles etc. Such a matchmaking beacon is important not only for information sharing about patients but to determine the

requirements for syntax, transmission of query, and transmission of phenotype information online. To bring together experts on these issues to agree on the framework would lead to greater adoption. This is an area of big impact if done right and taking into account appropriate privacy constraints.

The work of this proposed TF will be aligned with that of TGs of the GA4GH Clinical Working Group to define the exact relationship between the two groups and to explore possible actions in supporting the workshop. IRDiRC is in a strong position to lead and drive this forward.

Disease Ontology

While the work on phenotype ontology is advancing, annotations for diseases are incomplete and there are needs for better disease ontology research work, especially in the domain of neurobehaviour (e.g. in autism, mental retardation, and psychiatric manifestation of diseases), to improve the ability to understand these diseases. The utility of resources in this area is high but it may be hard to structure the ontology, impeded by the use of preferred terms in a given community, leading to clusters of different terms.

Main deliverables and actions

- ▶ Letter to inform members of the WG of IRDiRC restructuring
- ▶ Finalising and publication of the list of ICHPT terms
- ▶ Applications for “IRDiRC Recommended”
- ▶ Defining IRDiRC-GA4GH relationship
- ▶ Proposal of the TF to the Exec Comm