



# InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

## Terminology Cross-domain Task Force Research Statement

When research is carried out by a multidisciplinary and multicultural team that spans fifteen fields of inquiry and twenty countries, the precision and consistency of the terminology used in the course of the project is vital to the success of the research. Several terms that are key to this research project refer to different concepts in each disciplinary and/or cultural environment involved, while similar concepts are expressed by different terms. The Terminology Team is responsible for researching all terms proposed for official use by each research unit within InterPARES 2 and accepting or rejecting them on the basis of clarity, consistency with the other adopted terms, and validity in the various disciplinary and cultural contexts.

To reach its goal, the Terminology Cross-domain will be responsible for:

- Developing a register of terms and phrases to be selected from a corpus of InterPARES 1 and 2 documents and the UBC Project Glossary.
- Developing a Glossary that will provide logical or conceptual definitions of the words and phrases in the register as they are to be used for working purposes within IP2, to provide for consistency.
- Following the above activity, developing an “InterPARES 2 Dictionary” that will provide discipline-specific logical or conceptual definitions of the terms or phrases in the glossary (as well as additional terms and phrases that are not in the glossary) as they apply to the various disciplines.
  - The purpose of the discipline specific dictionary is allowing the findings of IP2 to be used by the communities that are affected by them, for example: to support the writing of guidelines. The discipline specific glossary will allow the findings of IP2, which are cast in IP specific terminology to be recast in the language of artistic, scientific and government disciplines for example.
- Developing the thesaurus (the list of terms and their relationships). The thesaurus will be referred to as the InterPARES 2 Thesaurus.

In light of all of the above product-related charges, the cross-domain research questions, which were formulated at the beginning of the research project, are to be considered a follow up to these additional questions, which must be addressed first. They are organised by expected product.

### **Register**

- What texts from the UBC Project, IP1 and 2 should be used to create the corpus of words that will be concorded?
- What terms and phrases in the corpus, to date, need to be carried forward as working terms and phrases? (Register is a collection of working terms)
- What working terms and phrases are candidates for working definitions?
- What format is the register going to be in?
- What rules and procedures should be followed to develop, revise, and add to the register?

### **Glossary**

- What definitions of the working terms and phrases exist already in the UBC Project, IP1 and IP2 official documents?
- Which among those definitions need to be revisited in light of the partial findings of IP2? (Iterative process)
- Which additional terms and phrases need to be defined in light of the partial findings of IP2?
- What are the form and components of a glossary entry?
- What rules and procedures should be followed to develop, revise, and add to the glossary?

### **Dictionary**

- What discipline specific texts should be included in the dictionary corpus in order to produce a register of terms and phrases?
- What terms or phrases should be included in the dictionary as headwords or head phrases in addition to those that are in the glossary?
- What are the form and components of a dictionary entry?
- What are the rules and procedures that should be followed to develop, revise, and add to the dictionary?

### **Thesaurus**

- What are the criteria for determining the syndetic (i.e., hierarchical and equivalence relationships – use for, broader term/narrower term) relationships between the terms and phrases for each concept expressed in the dictionary?
- What are the form and components of a thesaurus entry?
- What are the rules and procedures that should be followed to develop, revise, and add to the thesaurus?

### **Methodology:**

- Develop rules and procedures for generating all of the products.
- Develop criteria for deciding upon the terms and phrases for inclusion in the products.
- Generate the Register by means of concordance software operating on electronic copies of IP1 and 2 and UBC Project corpus.
- Create a bibliography of primary and secondary sources for IP2 terminology.

- Ask the RAs and the experts in specific disciplines to excerpt quotations from these sources for inclusion in the Glossary and the Dictionary.
- Generate products in the following order: Register, Glossary, Dictionary, and Thesaurus.
- Build an online Glossary/Dictionary/Thesaurus database for all researchers to view, provide feedback, provide suggestions, comments, criticisms, etc.
- Use feedback that comes in to revise the Glossary, Dictionary, and Thesaurus.
- Investigate and experiment with new computational methods of conceptual mapping. (e.g., topic maps)

**Timeline:**

- Develop database for terms, phrases, and definitions with the capability of capturing quotations and feedback from researchers by March 14, 2003.
- Draft Register due by March 14, 2003.
- Feedback from researchers on Draft Register due by March 28, 2003.
- Review annotated bibliographies with quotations by May 9, 2003. (Bibliographies must be received as soon as possible or as completed, but by April 14, 2003)
- Working Glossary due by June 13, 2003.
- All procedures and criteria for the Working Register and Working Glossary by June 13, 2003.
- Development of procedures and criteria for Dictionary during L.A. Plenary workshop. (September 22nd to 26th, 2003)
- Development of the First Working Dictionary by February 2004, for the L.A. Workshop.
- Developing procedures and criteria for Thesaurus during L.A. Workshop, February 2004.
- Development of the first Working Thesaurus by Vancouver Workshop in September 2004.
- Refining the instruments for final version from September 2004 to September 2006.