



InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

*International Research on Permanent Authentic
Records in Electronic Systems (InterPARES) 2:
Experiential, Interactive and Dynamic Records*

PART EIGHT

TERMINOLOGICAL INSTRUMENTS

Terminology Cross-domain Task Force Report

[including Appendix 22]

by

*Joseph T. Tennis, University of Washington
Randy Preston, The University of British Columbia*

- Status:** Final (public)
- Version:** Electronic
- Submission Date:** February 2007
- Publication Date:** 2008
- Project Unit:** Terminology Cross-domain Task Force
- URL:** http://www.interpares.org/display_file.cfm?doc=ip2_book_part_8_terminology_task_force.pdf
- How to Cite:** Joseph T. Tennis and Randy Preston, "Part Eight—Terminological Instruments: Terminology Cross-domain Task Force Report," [electronic version] in *International Research on Permanent Authentic Records in Electronic Systems (InterPARES) 2: Experiential, Interactive and Dynamic Records*, Luciana Duranti and Randy Preston, eds. (Padova, Italy: Associazione Nazionale Archivistica Italiana, 2008).
<http://www.interpares.org/display_file.cfm?doc=ip2_book_part_8_terminology_task_force.pdf>

Table of Contents

Introduction.....	1
Research Team.....	1
Team composition.....	1
Research mandate	2
Research Questions.....	3
Research Initiatives.....	3
Research Methodology	4
Terminology Database.....	4
Purpose.....	4
Scope.....	4
Structure.....	4
Statistics and structure of the terminological instruments	5
Dictionary	5
Glossary	6
Ontologies.....	6
Public access	7
Conclusions.....	7
Appendices	
Appendix 22: InterPARES 2 Project Ontologies.....	8

List of Figures

Figure 1. “Tool” entry in InterPARES 2 Project Dictionary	5
Figure 2. “Archives” entry in InterPARES 2 Project Glossary—illustrating a polysemous term.....	6
Figure 3. InterPARES 2 Project Ontology C - “Trustworthiness”	7

Introduction

When research is carried out by a multidisciplinary and multicultural team that spans more than one dozen fields of inquiry and twenty countries, the precision and consistency of the terminology used in the course of the research is vital to the success of the Project. Several terms that are key to this Project refer to different concepts in many of the disciplinary and/or cultural environments involved, while similar concepts are expressed by different terms. The term “record” is but one example. As used in the context of the InterPARES 2 research, a record is defined as “a document made or received in the course of a practical activity as an instrument or a by-product of such activity, and set aside for action or reference.”¹ This definition stands in stark contrast to that used by most of the other disciplines involved in the Project. For example, in Computer Sciences, a record is often defined as “an ordered set of fields, usually stored contiguously” or “a grouping of interrelated data elements forming the basic unit of a file,” while in the Arts a record refers to “any electronic, photographic or mechanical recording of music, singing, dialogue, sound effects or visual events, including CDs, DVDs, audio tapes, films, videos and the like.”²

The InterPARES 2 Terminology Cross-domain Task Force was responsible for researching and defining all terms proposed for official use by each research unit within the Project 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 this end, the Task Force developed a Terminology Database composed of three terminological instruments designed to be of service to each of the other research units within the Project and, by extension, to Archival Science. Over the course of the five-years of the Project, the terminology team collected words, definitions and phrases from extant documents, research tools and models, and through direct researcher submissions and discussions. From these raw materials, the team developed a systematic and pragmatic way of establishing a coherent view on the concepts involved in dynamic, experiential, and interactive records and systems in the arts, sciences and government.

Research Team

Team composition

Like each of the Project’s other cross-domains, the Terminology Cross-domain comprised researchers from a mixture of academic, archival and cultural heritage institutions, assisted by graduate research assistants from the universities of British Columbia and California, Los Angeles. During the first three and a half years of the Project, the Cross-domain was chaired by Jonathan Furner of the University of California, Los Angeles. In mid-2005, after Furner stepped down as chair, acting chairmanship was passed to Luciana Duranti before a new chair, Joe Tennis, was appointed in September 2005.

The following is a complete list of researchers and research assistants who participated in the Terminology Cross-domain Task Force at some point during the Project.

¹ InterPARES 2 Terminology Database. Available at http://www.interpares.org/ip2/ip2_terminology_db.cfm.

² Ibid.

Chairs:

Jonathan Furner	Jan 2001 - June 2005
Luciana Duranti	June - Sept 2005
Joe Tennis	Sept 2005 - Dec 2006

Researchers:

Barbara Craig	University of Toronto, Canada
Luciana Duranti	The University of British Columbia, Canada
Philip Eppard	University of Albany, State University of New York, USA
Jonathan Furner	University of California, Los Angeles, USA
Ian Lancashire	University of Toronto, Canada
Richard Pearse-Moses	Arizona State Library, USA
John Roeder	The University of British Columbia, Canada
Joe Tennis	University of Washington, USA
James Turner	Université de Montréal, Canada

Research Assistants:

David Boudinot	The University of British Columbia, Canada
Natalie Catto	The University of British Columbia, Canada
Naomi Cull	The University of Toronto, Canada
Kimberly Davison	The University of British Columbia, Canada
Shanna Fraser	The University of British Columbia, Canada
Jessica Glidewell	The University of British Columbia, Canada
Nadine Hafner	The University of British Columbia, Canada
Peggy Heger	The University of British Columbia, Canada
Eleanor Kleiber	The University of British Columbia, Canada
Tracey Krause	The University of British Columbia, Canada
Karen Langley	The University of British Columbia, Canada
Yvonne Loisselle	The University of British Columbia, Canada
Katherine Miller	The University of British Columbia, Canada
Emily O'Neill	The University of British Columbia, Canada
Carolyn Petrie	The University of British Columbia, Canada
Randy Preston	The University of British Columbia, Canada
Corinne Rogers	The University of British Columbia, Canada
Nadav Rouche	University of California, Los Angeles, USA
Anthea Seles	The University of British Columbia, Canada
Brian Trembath	The University of British Columbia, Canada
Melanie Wallace	The University of British Columbia, Canada
Catherine Yasui	The University of British Columbia, Canada

Research mandate

The InterPARES Project set out the following responsibility for the Terminology Research Team:

The Terminology Research Team will control the use of terms in all areas of the research. This research team will establish formal procedures for the proposal and adoption of specific terms, and meet in conjunction with International Team workshops to approve the official terms of the project and related definitions,

ensuring consistency among the various research units and keeping into account disciplinary and cultural differences.³

Research Questions

The InterPARES Project set out the following research questions for the Terminology Research Team:

- Is the term proposed specific to a field? If so, is its definition agreed upon in such field? If other definitions exist, how does the definition proposed relate to the others used? If the term is not specific to a field, is it a term in common usage or a neologism? How is its use justifiable in the context of the research?
- Is the term proposed used in other fields as well? If so, is its definition consistent across such fields? If not, what are the justifications for using one definition over another?
- Is the term used in several languages/traditions? If so, are the definitions consistent? If not, what are the justifications for using one definition over another?
- Is the term proposed consistent with the terms already used by the project? If so, does such inconsistency warrant a review of the already accepted terms in light of the new findings?
- Does the term express a concept that is already wholly or partially expressed by other already accepted terms or more appropriate terms?

Research Initiatives

To fulfil its mandate and address its research questions, the Terminology Team was tasked with developing a Register of terms and phrases to be selected from a corpus of InterPARES 1 and 2 documents and the earlier UBC Project Glossary.⁴ Using the terms in the Register, the team was then tasked with developing the following terminological instruments:

- 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 the Project, to provide for consistency.
- A 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.
- A Thesaurus (i.e., a list of terms and their relationships).⁵

Collectively, these instruments were to comprise what became known as the InterPARES 2 Terminology Database.

³ See Luciana Duranti (2001), "International Research on Permanent Authentic Records in Electronic Systems (InterPARES): Experiential, Interactive and Dynamic Records," SSHRC MCRI InterPARES 2 Project Proposal, 412-2001, 1.1-7. Available at http://www.interpares.org/display_file.cfm?doc=ip2_detailed_proposal.pdf.

⁴ See <http://www.interpares.org/UBCProject/gloss.htm>.

⁵ This initiative began as a class project when a group of students at the University of California at Los Angeles, who had been working on a re-design of the US-InterPARES Web site, were encouraged by the then Terminology Team chair, Jonathan Furner, to build a thesaurus as part of the Web site redesign effort. The students began with the InterPARES Glossary and built a faceted, hierarchical structure on top, using the TCS-8 thesaurus construction tool available from WebChoir (<http://www.webchoir.com/>). Although never completed, a draft version of the thesaurus, which contains approximately 1,200 terms, is available online at <http://polaris.gseis.ucla.edu/jfurner/IQ2/home.htm>.

Research Methodology

The typical approach for populating the Terminology Database was for the terminology team to receive proposed terms from each of the Project's research units, together with proposed definitions. The terminology team's task was to then research the term according to the research questions listed above, by examining past InterPARES and UBC Project research documents, as well as the relevant dictionaries, glossaries and literature of the fields and countries represented in the Project. The team also consulted general dictionaries, where appropriate. On the basis of the result of such research, the team either: (1) accepted a term with its submitted definition, (2) returned it to the relevant research unit with proposed changes to the definition and/or a proposed alternate term or (3) rejected the term.

Terminology Database

Purpose

The original purpose of the InterPARES 2 Terminology Database was to support the Project's researchers in understanding concepts across disciplines that have come to address issues of preserving dynamic, experiential and interactive authentic digital records in electronic systems. Archival Science, Information Science, Computer Science, Geomatics, Music, Film, Dance, Law and several other disciplines all brought expertise and terminology to this Project. The Terminology Database facilitated communication and research among researchers in different disciplines and across cultural boundaries throughout the course of the Project by defining terms and relationships among terms used in the research. By extension, the Terminology Database now stands as a significant contribution to the understanding of records in dynamic, interactive and experiential systems in the artistic, scientific and governmental sectors.

Scope

To prepare the terms and the definitions for the Terminology Database, the team consulted resources relevant to each of the fields and disciplines represented in the Project. These included Arts (Film, Music, Dance), Sciences (Geomatics, Physics, Astronomy, Archaeology), Government (Law, Policy, E-Government) and Archival, Information and Computer Sciences. Terminology developed and used in InterPARES 1 and in the earlier UBC Project was carried forward to the InterPARES 2 Terminology Database. In this sense, the InterPARES 2 Project built upon its past knowledge, making it relevant to the contemporary environment.

Structure

There are two main components to the Terminology Database. They are (1) the Glossary and (2) the Dictionary. The Glossary is the authoritative list of terms and definitions that were core to the researchers' understanding of the evolving records creation, recordkeeping and records preservation environments. As the authoritative list, the Glossary provided the approved terms and definitions used throughout the InterPARES 2 research in both working documents and published documents. Because many, if not most, of the terms that were approved have multiple cross-disciplinary and/or cross-cultural definitions, it was important to account for plurality of meaning, which is precisely what the Dictionary is intended to do. The Dictionary contains all

the terms from the Glossary, but, unlike the Glossary, the Dictionary provides multiple definitions for a single term gleaned from multiple disciplines. Each of these definitions are cited as coming from a particular discipline (e.g., Science, Archives or Arts), and from a particular source (e.g., Art and Architecture Thesaurus, Dictionary of Computing, or A Glossary of Archival and Records Terminology). Sources are of two kinds: dictionaries in the field or research documents from InterPARES 1 or 2. In short, the Dictionary is a tool used to facilitate interdisciplinary communication among members of the various communities that have a stake in the use of the terms used in the Project; for example, to support the writing of guidelines. The discipline-specific definitions of terms in the Dictionary has allowed the findings of the Project, many of which are cast in InterPARES-specific terminology, to be recast in the language of artistic, scientific and government disciplines. Thus, for example, by using this tool, Project researchers were able to see how Archival Science deploys terminology compared to researchers and practitioners in Computer Science, Library and Information Science, the Arts, etc., as demonstrated above in the introduction for the term “record.”

A third terminological instrument, the Ontologies, was developed to identify explicit relationships among concepts. This instrument is particularly useful for communicating the nuances of Diplomats in the dynamic, experiential and interactive environment.

All three of these instruments were drawn from a Register of terms gathered over the course of the Project. This Register served as a holding place for terms and phrases, and allowed researchers to discuss, comment on and modify submissions. The Register and the terminological instruments were housed in the online Terminology Database. The Database provides searching, display and file downloads, making it easy for users to navigate through the terminological instruments.

Statistics and structure of the terminological instruments

Dictionary

The Dictionary is the largest of the terminological instruments. It contains more than 900 terms, most of which contain multiple definitions. As shown in Figure 1, the entries in the Dictionary follow a standard format: term, part of speech, definition, disciplinary classification and citation. As well, it is important to note that, in cases where a Dictionary entry also appears in the Glossary, the Glossary definition is always the first of any multiple definitions listed for the entry in the Dictionary.

tool

n., Information technology, or other equipment or supply used to manage records at any time during their lifecycle. [Archives - *Manage Chain of Preservation*, Page: n.p.]

n., A program used primarily to create, manipulate, modify, or analyse other programs, such as a compiler or an editor or a cross-referencing program. Opposite: app, operating system. [Computer and Information Sciences - *Free Online Dictionary of Computing (FOLDOC)*]

n., Any existing physical object that is in some way fashioned or altered by humans and employed for a specific task or purpose. [Sciences - *Encyclopedic Dictionary of Archaeology*]

Figure 1. “Tool” entry in InterPARES 2 Project Dictionary

Glossary

The Glossary is smaller than the Dictionary because it includes the only meaning in which each term is used within InterPARES 2 documents. The Glossary terms with their definitions are the key to the communication of the findings of InterPARES 2. There are more than 450 terms in the Glossary (excluding terms and phrases that are specific to the Chain of Preservation and Business-driven Recordkeeping models). The format is similar to the Dictionary, but, as already stated, with only a single definition for each term. Exceptions arise when there is a single term for many concepts, such as in the example provided in Figure 2.

archives

n., [records] The whole of the documents made and received by a juridical or physical person or organization in the conduct of affairs, and preserved. [Archives - *School of Library, Archival and Information Studies (SLAIS) Glossary*]

n., [institution] An agency or institution responsible for the preservation and communication of records selected for permanent preservation. [Archives - *The InterPARES 1 Project Glossary*]

n., [place] A place where records selected for permanent preservation are kept. [Archives - *The InterPARES 1 Project Glossary* , Page: 356]

Figure 2. “Archives” entry in InterPARES 2 Project Glossary—illustrating a polysemous term

A key support mechanism for both the Glossary and Dictionary is the InterPARES 2 Bibliography. Where possible, the Project rooted its definitions in extant literature and cited the source for each definition that appears in the Terminology Database. This truly provides context for the Project’s definitions because it stitches the use of the words and concepts as understood by the Project’s researchers with the same words and concepts used in other texts by researchers and practitioners outside of the Project.

Ontologies

To illustrate explicit relationships among terms in the Glossary, the Terminology Cross-domain constructed several Ontologies. Within the context of the Project, the Ontologies were used to illustrate the relationship among Diplomatic terms. The terminology team constructed three Ontologies: (1) Archives and Records, (2) Status of Transmission and (3) Trustworthiness.⁶ All of these Diplomatic concepts are core to the Project’s understanding of records in dynamic, experiential and interactive systems and are made clearer through specific explication in the Ontologies. Figure 3 provides an example of an ontology by illustrating the component parts of Trustworthiness and their relationships.

⁶ See Appendix 22.

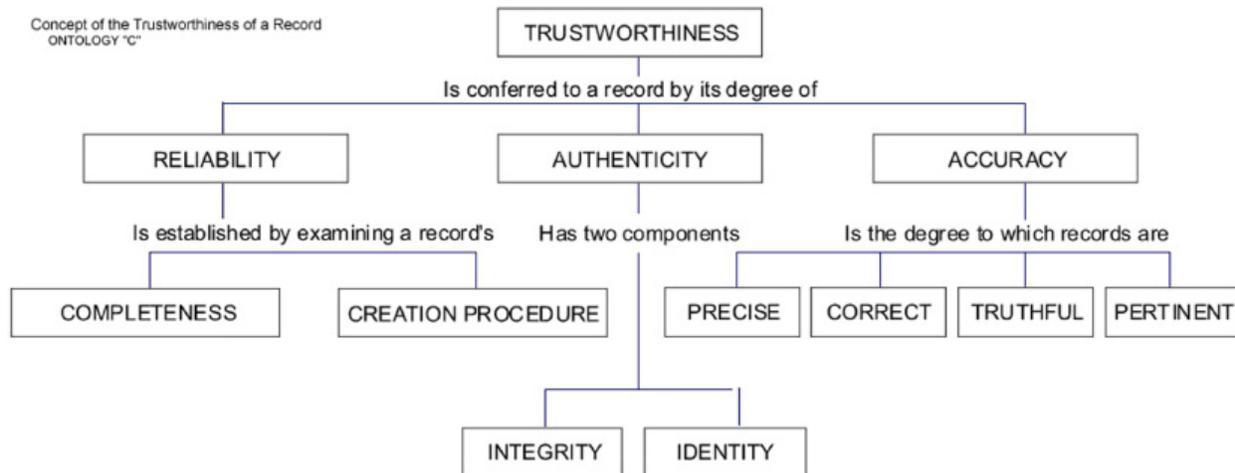


Figure 3. InterPARES 2 Project Ontology C - “Trustworthiness”

Public access

At the close of the InterPARES 2 Project, the Terminology Database was “frozen” and is available to the general public via the InterPARES Web site.⁷ The Dictionary and Glossary are searchable and downloadable in PDF format, and the Ontologies are available in PDF format.

Conclusions

The Terminology Cross-domain Task Force acted as a service research team to the InterPARES 2 Project. It is hoped that it fulfilled the same role to Archival Science. In the process of constructing the three terminological instruments discussed above, the Project made great steps in advancing the general understanding of Archival concepts in the contemporary environment and expanded the existing records- and preservation-related vocabulary to include new ways of thinking about age-old problems.

Terminology work is not without its challenges. The challenges of terminological work—on any scale—are challenges of time and socio-political negotiation. As knowledge develops over time new terms surface, old terms are reinterpreted, forgotten, ossified in their original context, or brought forward into a new and vibrant scholarly discourse. Archival terminology is rich in expressiveness and history. As a service research team, the Terminology Cross-domain aimed at keeping terminological tradition, innovation and integration a positive and productive venture for the Project researchers and for Archival Science.

⁷ See http://www.interpares.org/ip2/ip2_terminology_db.cfm.

Appendix 22

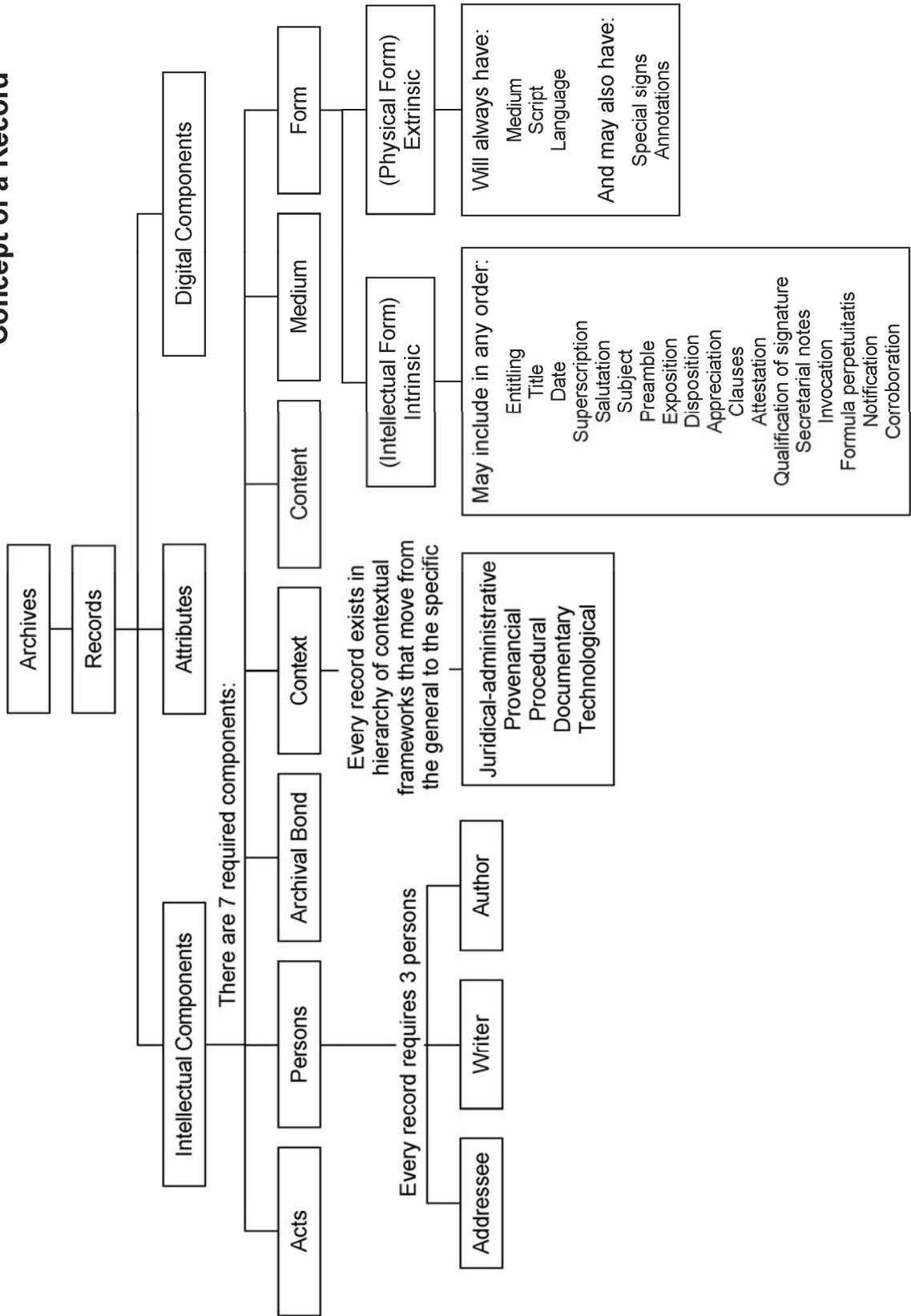
InterPARES 2 Project Ontologies

Ontology A: Concept of a Record

Ontology B: Concept of the Status of Transmission of a Record

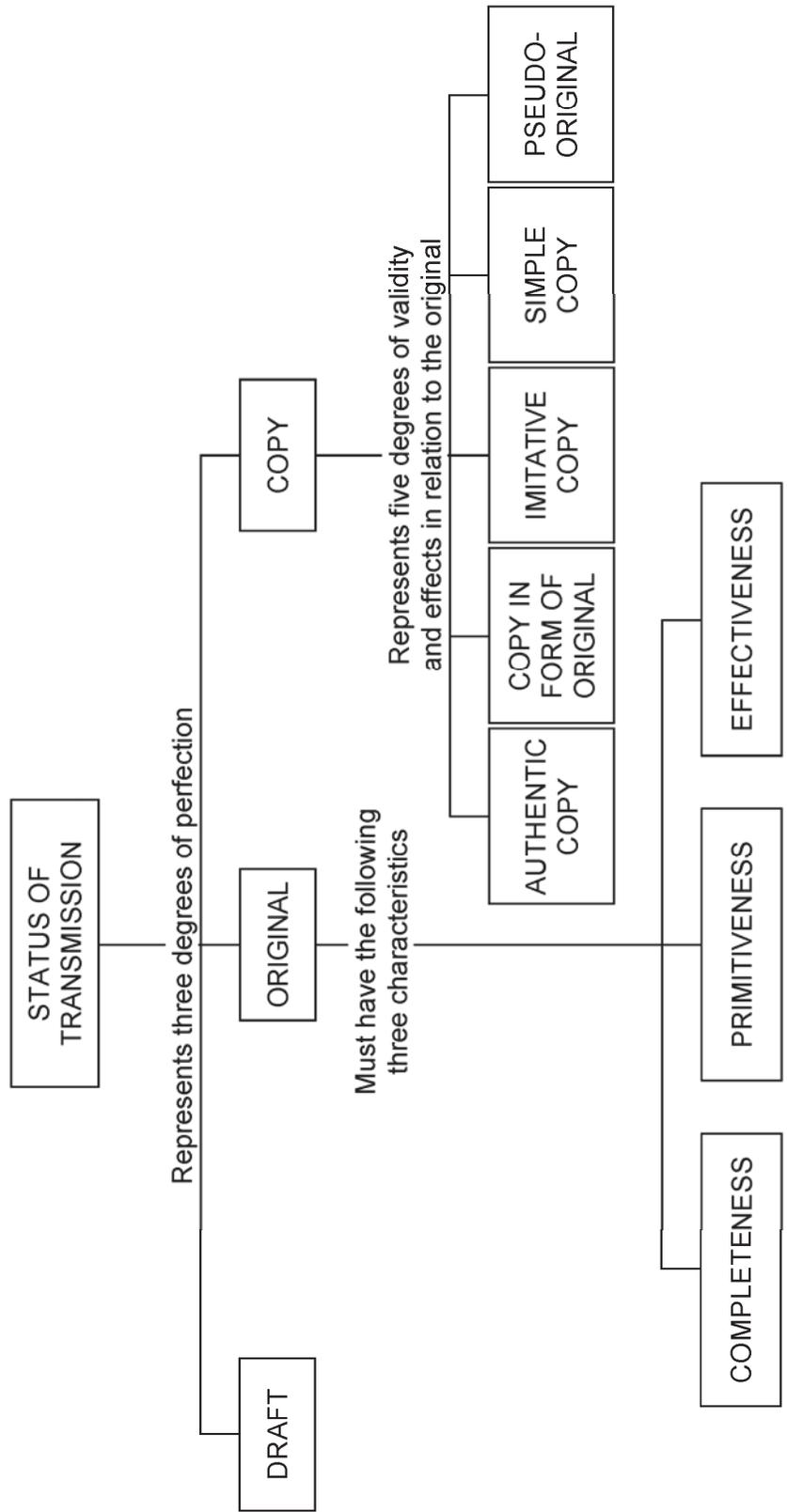
Ontology C: Trustworthiness of a Record

**ONTOLOGY A:
Concept of a Record**



**ONTOLOGY B:
 Concept of the Status of Transmission of a Record**

A document's degree of authority depends on its





ONTOLOGY C: Trustworthiness of a Record

