



InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

Title: Case Study 12 Final Report: Antarctic Treaty Searchable Database

Status: Final (public)

Version: 1.0

Submission Date: March 2005

Release Date: September 2007

Author: The InterPARES 2 Project

Writer(s): Paul Arthur Berkman
Bren School of Environmental Science and Management,
University of California, Santa Barbara
George James Morgan, III, EvREsearch, Ltd.
Reagan Moore, San Diego Supercomputer Center
Richard Marciano, San Diego Supercomputer Center
Jim Suderman, Archives of Ontario
Babak Hamidzadeh, Library of Congress
Hans Hofman, National Archives of The Netherlands

Project Unit: Focus 3

URL: http://www.interpares.org/display_file.cfm?doc=ip2_cs12_final_report.pdf

Table of Contents

List of Figures	ii
A. Overview	1
B. Statement of Methodology.....	4
Implementation of the Antarctic Treaty Searchable Database	4
“Dynamic, Interactive and Experiential Components”	4
Outreach Assessments	5
C. Description of Context.....	6
Provenancial Context	6
Juridical-administrative Context	10
Procedural Context.....	11
Documentary Context	14
Technological Context.....	15
D. Addressing the 23 Core Research Questions.....	19
Conclusion	41
E. References Cited.....	43
F. Glossary of Terms	44
G. IDEF0 Activity Model	50
Appendix 1: Abstracts	59

List of Figures

Figure 1.	Organizational matrix of the <i>International Research on Permanent Authentic Records in Electronic Systems (InterPARES 2)</i> program.....	1
Figure 2.	Comprehensive hierarchy showing the relationships within and between years for the 120 adopted measures in the Antarctic Treaty System that refer to ‘information exchange.’.....	7
Figure 3.	Comprehensive hierarchy showing the relationships within and between years for the 55 adopted measures in the Antarctic Treaty System that refer to “secretariat.”.....	8
Figure 4.	A generalized Activity-Flow-Diagram to illustrate the processes to create the <i>Antarctic Treaty Searchable Database</i> or other digital records with the <i>Digital Integration System™ (DIGIN®)</i> from EvREsearch LTD.	13
Figure 5.	A Data-Flow-Diagram to illustrate the activities along with specific data elements and data stores to implement <i>Antarctic Treaty Searchable Database</i> with integrated access to the policy documents from the Antarctic Treaty System.	14
Figure 6.	The <i>Digital Integration System™ (DIGIN®)</i> from EvREsearch LTD.....	16
Figure 7.	Cumulative profile of Antarctic Treaty “ <i>measures</i> ” with search terms that were comprehensively integrated over time with the <i>Antarctic Treaty Searchable Database</i> , as expressed by hierarchal displays.	18
Figure 8.	Illustration of the exponentially increasing number of possible permutations or relationships among ‘N’ objects.	20
Figure 9.	A model of the exponentially increasing volume of metadata by simply doubling the number of information granules (data).	25
Figure 10.	Cumulative profile of relationships among Antarctic Treaty “ <i>measures</i> ” with 2, 3 or 4 of the search terms from Figure 6 that were comprehensively integrated over time with the <i>Antarctic Treaty Searchable Database</i> , as expressed in hierarchal displays	32
Figure 11.	Thresholds in the preservation and dissemination of written information in our civilization.....	41

A. Overview

The *Antarctic Treaty Searchable Database* was approved as a case study for the InterPARES 2 Project at its international meeting in Rome, Italy, in November 2002. The *International Research on Permanent Authentic Records in Electronic Systems* project involves national archives and representatives from sixteen countries who are investigating the “*dynamic, interactive and experiential components*” of archival activities that are unique to digital environments. The *Antarctic Treaty Searchable Database* and the other 25 case studies in the InterPARES 2 Project have been tasked with answering questions about the provenancial, juridical, procedural, documentary and technological components of their records. Results from these case studies are being compiled and synthesized into a set of recommendations regarding archival best practices in digital environments.

The InterPARES 2 Project has been designed from 2001–2006 to address a matrix of elements that are integral to the creation, management and preservation of electronic records that are incorporated into archives (Fig. 1). The *Antarctic Treaty Searchable Database* is aligned with Focus 3 on *Government Activities* as well as Domain 1 on *Records Creation and Maintenance* and Domain 2 on *Authenticity, Accuracy and Reliability*. This international case study also contributes to the *Policy* and *Description* Cross-domains.

	FOCUS 1 Artistic Activities	FOCUS 2 Scientific Activities	FOCUS 3 Government Activities
DOMAIN 1 Records Creation and Maintenance	Working Group 1,1	Working Group 1,2	Working Group 1,3
DOMAIN 2 Authenticity, Accuracy and Reliability	Working Group 2,1	Working Group 2,2	Working Group 2,1
DOMAIN 3 Methods of Appraisal and Preservation	Working Group 3,1	Working Group 3,2	Working Group 3,3
Terminology Cross-Domain			
Policy Cross-Domain			
Description Cross-Domain			
Modeling Cross-Domain			

Figure 1. Organizational matrix of the *International Research on Permanent Authentic Records in Electronic Systems* (InterPARES 2) program (<http://www.interpares.org>).

As a case study, the *Antarctic Treaty Searchable Database* involves a dynamic and interactive record of all recommendations, conventions, measures, decisions, resolutions, annexes, appendices, tables, figures and protocols that have been formally adopted by the Antarctic Treaty

Consultative Parties starting with *1959 Antarctic Treaty* to the present. Overall objectives of the *Antarctic Treaty Searchable Database* are to:

- a. facilitate comprehensive, objective, user-defined discovery of relationships between the adopted measures;
- b. provide flexible framework for updating, integrating and preserving the record with new measures;
- c. provide a cost-effective and accessible system that maintains accuracy, authenticity and reliability of the digital record over time;
- d. identify specifications and policies to create, preserve and enhance the functionality of authentic digital records;
- e. enhance the interdisciplinary applications of the Antarctic Treaty System and other policy systems for government, education, industry and public purposes; and,
- f. demonstrate strategies to objectively integrate digital information in a manner that stimulates inquiry, facilitates interpretation and opens doors for knowledge discovery.

With funding from the National Science Foundation and materials from the United States Department of State, the *Antarctic Treaty Searchable Database* was first produced by Paul Berkman and George Morgan in early 2000. Shortly afterward, this database of public-domain records was introduced by the United States Department of State at the 23rd Antarctic Treaty Consultative Meeting (ATCM) in Peru. The *Antarctic Treaty Searchable Database*, which is available online¹ as well as on webCDserverSM, now is in its 5th Edition with all of the measures that have been adopted by the Antarctic Treaty Consultative Parties through 2004.

While the *Antarctic Treaty Searchable Database* was intended originally as a supplement for university courses on Antarctic science and policy, it soon became linked to:

- international government institutions (e.g., Antarctic Treaty Secretariat);
- government agencies (e.g., Australian Antarctic Division, Environment Canada; and the United States Library of Congress);
- non-governmental organizations (e.g., Scientific Committee on Antarctic Research and Antarctic Southern Ocean Coalition);
- business (e.g., International Association of Antarctic Tour Operators); and
- academic institutions (e.g., George Washington University; Universität Freiburg, Oxford University and the University of California Santa Barbara).

¹ See <http://aspire.tierit.com/>.

Moreover, the *Antarctic Treaty Searchable Database* was linked to the international Web sites for the 45 nations contributing to the 24th (2001) and 25th (2002) ATCM in St. Petersburg, Russia, and Warsaw, Poland, respectively.

Nine key rules were developed by Paul Berkman to identify, incorporate, organize, integrate and display the Antarctic Treaty measures with the *Antarctic Treaty Searchable Database*. These rules and their implementation are described in relation to the: (a) juridical; (b) provenancial; (c) procedural; (d) documentary; and (e) technological activities of an archive.

Activity-flow and data-flow diagrams have been constructed to describe the ongoing implementation of the *Antarctic Treaty Searchable Database* with the *Digital Integration System*TM (*DIGIN*[®]) from EvREsearch LTD. *DIGIN*[®] is an automated technology that objectively integrates digital record entities without markup, metadata or databases. Unlike lists of information, for every query, the *Antarctic Treaty Searchable Database* dynamically generates hierarchal displays that objectively describe conceptual relationships within and between the Antarctic policy documents. Functionalities of these dynamic hierarchies are further contrasted with the digital version of the *Antarctic Treaty Handbook*² that was most recently updated in 2002 by the Department of State in the United States (as depository government for the *1959 Antarctic Treaty*) with its twelve ‘pdf’ files that only can be searched sequentially one file at a time like any static paper document.

In addition to being the first searchable database of Antarctic Treaty documents ever produced, the *Antarctic Treaty Searchable Database* has nearly 750 digital record entities and is the most comprehensive source for integrating policy documents from the Antarctic Treaty System. Overall, the *Antarctic Treaty Searchable Database* is considered in this report with the goal of contributing to the practical implementation of the Antarctic Treaty Secretariat, which was approved in 2001 under Decision XXIV-1 and which began operation in Argentina with its first Executive Secretary in September 2004.

² Available at <http://www.state.gov/g/oes/rls/rpts/ant/>.

B. Statement of Methodology

Implementation of the Antarctic Treaty Searchable Database

The *Antarctic Treaty Searchable Database*³ is described in relation to its digital and hardcopy sources as well as the information technology procedures to:

- create the granularity of the policy documents;
- create categorical tags for each of the resulting policy granules; and
- index the collection of tagged policy granules.

These information technology procedures are described in relation to the *Digital Integration System*TM (*DIGIN*[®]) from EvREsearch LTD, which is based on their patented *Information Management, Retrieval and Display Systems and Associated Methods*. Data and activity flow diagrams (Bobak 1997) have been constructed to describe the sources, processes, stores and overall flow of the data to implement the *Antarctic Treaty Searchable Database*.

“Dynamic, Interactive and Experiential Components”

Implementation of the *Antarctic Treaty Searchable Database* is contrasted with conventional information management methodologies with metadata, mark-up and databases. These comparisons are based on:

- strategies to compile digital records and integrate digital record entities;
- comprehensive and objective relational displays; and,
- quantitative analyses of the relational displays.

Implementation of the *Antarctic Treaty Searchable Database* also is compared with the *Antarctic Treaty Handbook. 9th Edition* from the United States Department of State in 2002, which has 12 ‘pdf’ files that are available over the Internet.⁴ Technological applications are interpreted further in relation to the following aspects of the digital records:

- authenticity;
- accuracy;
- fixity;
- reproducibility;
- interoperability; and
- persistence.

As a composite digital record, the *Antarctic Treaty Searchable Database* is further considered in relation to digital record groups, digital record series and digital record entities that are implemented in archival practices.

³ Available at <http://aspire.tierit.com/>.

⁴ See <http://www.state.gov/g/oes/rls/rpts/ant/>.

Outreach Assessments

The *Antarctic Treaty Searchable Database* involves interdisciplinary applications for government, education, industry and the general public. To address the core questions as well as the questions from the domains and cross-domains, information was collected from:

- e-mail correspondences with government officials in the United States and abroad;
- meetings with national representatives from various foreign offices;
- Web site linkages to site the *Antarctic Treaty Searchable Database* by various entities;
- actual experience from implementing the database;
- questions that were raised after presentations at various conferences;
- interaction with personnel from the National Science Foundation, United States Department of State Marine Mammal Commission and National Archives and Records Administration; and,
- comments about the webCDserver of the *Antarctic Treaty Searchable Database* that was included in the textbook on *Science into Policy Global Lessons from Antarctica* (Berkman 2002).

These data interpretations are supported by objectively-generated data that are reflected by the tables and figures in this report. In addition, integrated searches with the *Antarctic Treaty Searchable Database* itself were used to address the 23 core research questions of the InterPARES 2 Project, as appropriate.

C. Description of Context

Provenancial Context (the creating body, its mandate, structure, and functions)

Originally intended as a supplement for a university course on Antarctic science and policy (Berkman 2002), the *Antarctic Treaty Searchable Database* represents an evolving digital record that has been created and maintained to benefit the diverse community of Antarctic stakeholders. The history of the *Antarctic Treaty Searchable Database* goes back to 1998, when the United States Department of State was contacted by Paul Berkman about availability of their *Antarctic Treaty Handbook. 8th Edition (1994)* in digital form. This query was prompted because the *Antarctic Treaty Handbook* had become unwieldy for teaching purposes with over 1000 pages and information management was rapidly moving toward digital forms. The response of the Department of State, at that time, was that the United States satisfied its role as depository government by producing hardcopy versions of the *Antarctic Treaty Handbook*. Without forethought, Paul Berkman responded that he was going to develop a searchable database of the Antarctic Treaty documents—and so began the odyssey.

The *Antarctic Treaty Searchable Database*, which is available online as well as on webCDserver, is in its 5th Edition with all of the “*measures that have been adopted in furtherance of the principals and objectives of the Treaty*” from 1959 through 2004. In addition to being the first searchable digital library of Antarctic Treaty documents ever produced, the *Antarctic Treaty Searchable Database* now has nearly 750 individual policy granules and remains as the most comprehensive source for automatically integrating information from the Antarctic Treaty System.

Two years after the first version of the *Antarctic Treaty Searchable Database* was produced and introduced at the 23rd Antarctic Treaty Consultative Meeting (ATCM) in Lima in 1999, the Antarctic Treaty Consultative Parties fundamentally changed “*information exchange*” in the Antarctic Treaty System by adopting Decision XXIV-1 at the 24th ATCM:

That the Antarctic Treaty Secretariat shall be established in Buenos Aires following the development of necessary modalities and agreements, which the Parties shall urgently pursue.

As these international negotiations regarding the Antarctic Treaty Secretariat were underway, the *Antarctic Treaty Searchable Database* was linked to the Web sites for the 24th and 25th ATCM in St. Petersburg and Warsaw, respectively.

A principal responsibility of the Antarctic Treaty Secretariat is the implementation of a “*central information exchange web site to be hosted by Argentina.*” The overall context of information exchange and “secretariat” in the Antarctic Treaty System are shown in Figures 2 and 3, respectively.

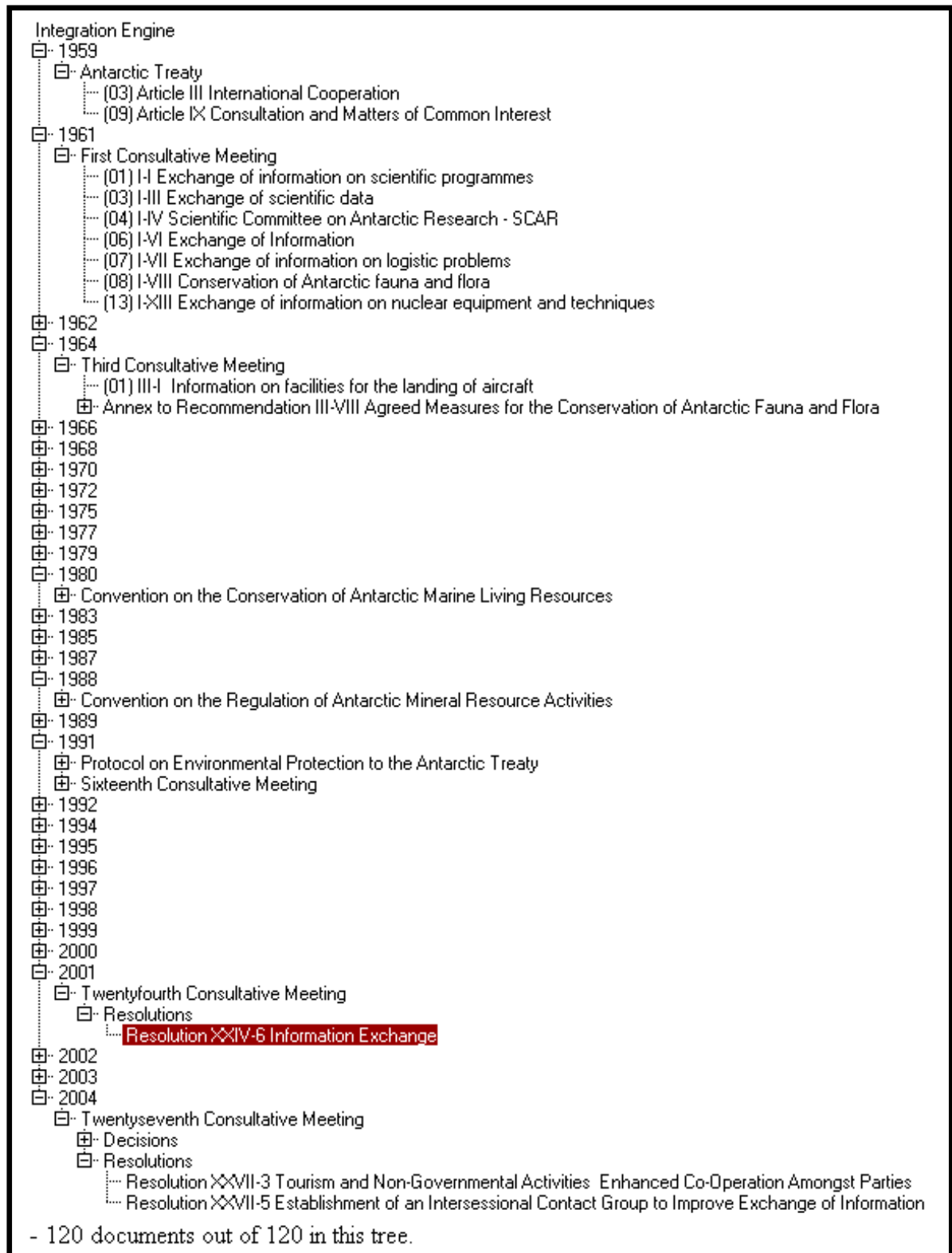


Figure 2. Comprehensive hierarchy showing the relationships within and between years for the 120 adopted measures in the Antarctic Treaty System that refer to ‘information exchange.’ It should be noted that measures dealing with information exchange have been adopted at every Antarctic Treaty Consultative Meeting following the ratification of the *1959 Antarctic Treaty*. This hierarchy was dynamically generated from the *Antarctic Treaty Searchable Database* (<http://aspire.tierit.com/>).

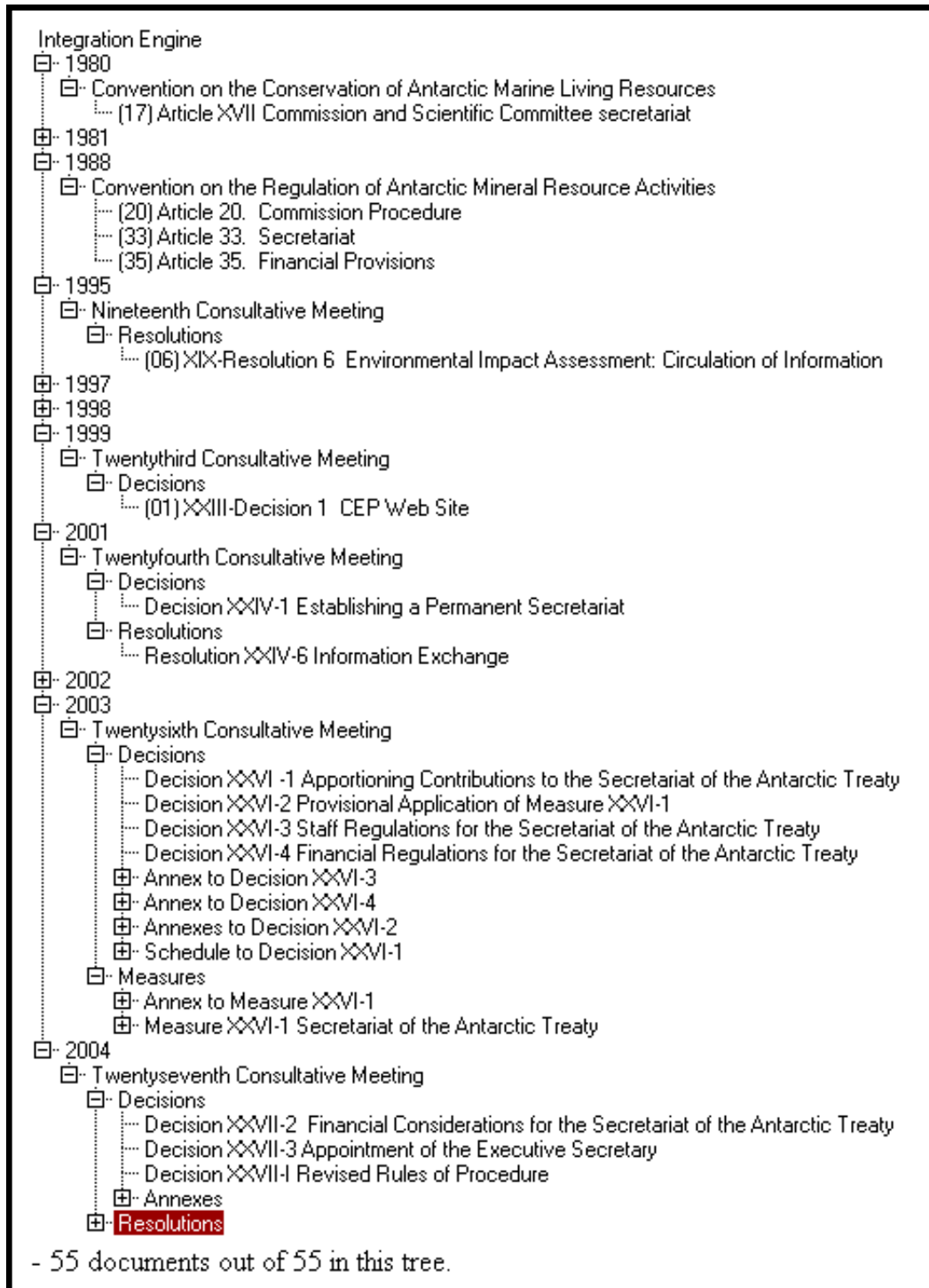


Figure 3. Comprehensive hierarchy showing the relationships within and between years for the 55 adopted measures in the Antarctic Treaty System that refer to “secretariat.” This hierarchy was dynamically generated from the *Antarctic Treaty Searchable Database* (<http://aspire.tierit.com/>).

With adoption of the Antarctic Treaty Secretariat at the 24th ATCM, information-exchange activities by the United States (as depository government of the *1959 Antarctic Treaty*) began to change. In 2002, the Department of State produced the *Antarctic Treaty Handbook. 9th Edition*. In addition, the Department of State also produced twelve ‘pdf’ files (Box 1) that were copied onto a CD-ROM and their Web site (<http://www.state.gov/g/oes/rls/rpts/ant/>) to mirror the organization and functionality of the hardcopy version of the *Antarctic Treaty Handbook. 9th Edition* (2002).

BOX 1

ANTARCTIC TREATY DOCUMENTS (IN ‘PDF’ FORMAT) ON THE UNITED STATES DEPARTMENT OF STATE WEBSITE

(<http://www.state.gov/g/oes/rls/rpts/ant/>)

[Chapter I -- Foreword to the Ninth Edition](#)
[Chapter II -- The Antarctic Treaty System: Introduction](#)
[Chapter III -- Operation of the Antarctic Treaty System](#)
[Chapter IV -- Inspections Under Article VII of the Treaty](#)
[Chapter V -- Exchanges of Information, Including Data Management](#)
[Chapter VI -- Scientific Cooperation](#)
[Chapter VII -- Logistical and Operational Issues](#)
[Chapter VIII -- Tourism and Other Non-governmental Activities](#)
[Chapter IX -- Conservation of Antarctic Seals: CCAS](#)
[Chapter X -- Conservation of Antarctic Marine Living Resources: CCAMLR](#)
[Chapter XI -- Regulation of Antarctic Mineral Resource Activities: CRAMRA](#)
[Chapter XII -- Protection of the Antarctic Environment](#)

From 1998 through 2004, the *Antarctic Treaty Searchable Database* has been implemented by EvREsearch LTD with partial support from the National Science Foundation. Initially, in 1998-99, funding for the *Antarctic Treaty Searchable Database* came from the Division of Undergraduate Education and Office of Polar Programs at the National Science Foundation. Since 2003, the *Antarctic Treaty Searchable Database* has been supported by the National Science Digital Library program at the National Science Foundation as part of a larger project with the Marine Mammal Commission on international environmental and ecosystem policy documents. In addition, the *Antarctic Treaty Searchable Database* was approved by InterPARES 2 as a case study at their meeting in Rome in 2002.

Although the *Antarctic Treaty Searchable Database* has been widely supported and applied as a comprehensive digital record of the “*measures that have been adopted in furtherance of the principals and objectives of the Treaty*,” it preceded and has yet to be designed around any formal government mandate. Informally, the *Antarctic Treaty Searchable Database* has been designed to facilitate knowledge discovery about the policies and strategies that promote “*international cooperation*” and the “*use of Antarctica for peaceful purposes only*” as stated in the *Preamble* of the *1959 Antarctic Treaty* (Berkman 2002).

Juridical-administrative Context (the legal and organizational system)

International creation of the policy documents for the Antarctic Treaty region, “*south of 60° south latitude,*” are mandated under Article IX of the *1959 Antarctic Treaty*:

1. Representatives of the Contracting Parties named in the preamble to the present Treaty shall meet at the City of Canberra within two months after the date of entry into force of the Treaty, and thereafter at suitable intervals and places, for the purpose of exchanging information, consulting together on matters of common interest pertaining to Antarctica, and formulating and considering, and recommending to their Governments, measures in furtherance of the principles and objectives of the Treaty, including measures regarding:

- a. use of Antarctica for peaceful purposes only;*
- b. facilitation of scientific research in Antarctica;*
- c. facilitation of international scientific cooperation in Antarctica;*
- d. facilitation of the exercise of the rights of inspection provided for in Article VII of the Treaty;*
- e. questions relating to the exercise of jurisdiction in Antarctica;*
- f. preservation and conservation of living resources in Antarctica.*

The international policy measures that have been approved by the Antarctic Treaty Consultative Parties between 1959 and the present were incorporated in the digital record of the *Antarctic Treaty Searchable Database*. These measures include the *1959 Antarctic Treaty* as well as the recommendations, measures, decisions, resolutions, annexes, appendices, attachments, tables and figures. In addition, the principal Conventions, Protocol and all of their parts (Table 1) that have emanated from the Antarctic Treaty Consultative Meetings have been included in the *Antarctic Treaty Searchable Database*.

The *Antarctic Treaty Searchable Database*, which was created by Paul Berkman and George Morgan with funding by the National Science Foundation in collaboration with the United States Department of State, explicitly states that:

This searchable database is being developed to enhance public access to the Antarctic Treaty documents and any errors or omissions are included from the original versions received from the United States Department of State, Marine Mammal Commission and National Science Foundation. There is no warranty, expressed or implied, as to the accuracy or completeness of the furnished data or the resulting searchable database displays.

In addition, there is an End-Users License Agreement that each user must accept before utilizing the database. Among its various components, this agreement further identifies that the licensor (i.e., EvREsearch LTD) disclaims “*all warranties of title, merchantability or fitness for a particular purpose*” with regard to *Antarctic Treaty Searchable Database*.

Table 1. Antarctic Resource Management Regimes

Antarctic Document	Year Signed	Year Ratified	Depository Government	Associated Institutions
Antarctic Treaty	1959	1961	United States	<i>Specialized agencies of the United Nations and other international organizations having a scientific or technical interest in Antarctica</i>
Seals Convention ¹	1972	1978	United Kingdom	Scientific Committee on Antarctic Research
Living Resources Convention ²	1980	1984	Australia	CCAMLR Commission, Scientific Committee, Secretariat and Arbitral Tribunal
Mineral Resources Convention ³	1988	not ratified	New Zealand	CRAMRA Commission, Advisory Committee, Regulatory Committees, Secretariat and Arbitral Tribunal
Environmental Protocol ⁴	1991	1998	United States	PROTOCOL Committee on Environmental Protection and Arbitral Tribunal along with: <i>Annex I: Environmental Impact Assessment;</i> <i>Annex II: Conservation of Antarctic Fauna and Flora;</i> <i>Annex III: Waste Disposal and Management;</i> <i>Annex IV: Prevention of Marine Pollution</i> <i>Annex V: Area Protection and Management</i>

1 *Convention on the Conservation of Antarctic Seals (CCAS)*

2 *Convention on the Conservation of Antarctic Marine Living Resources (CCAMLR)*

3 *Convention on the Regulation of Antarctic Mineral Resource Activities (CRAMRA)*

4 *Protocol on Environmental Protection to the Antarctic Treaty (PROTOCOL)*

Procedural Context (the business procedure to create the digital record)

The initial information resource for the *Antarctic Treaty Searchable Database* was the *Antarctic Treaty Handbook. 8th Edition* that was published by the United States Department of State in 1994. The *Antarctic Treaty Handbook* included most of the policy documents that had been adopted by the Antarctic Treaty Consultative Parties. These policy documents were organized into topical sections, which each had “introductory notes” added by the United States Department of State along with “extracts” from the reports of the Antarctic Treaty Consultative Meetings (ATCM). In addition to the *Antarctic Treaty Handbook*, contents for the *Antarctic Treaty Searchable Database* came from other sources or ‘data elements’ that were in either digital or hardcopy formats.

Based on the characteristics of the *Antarctic Treaty Handbook.8th Edition* (1994), the following rules were used to establish the parameters for compiling the contents of the *Antarctic Treaty Searchable Database*:

- Rule 1: Include only the “measures” that were adopted by the Antarctic Treaty Consultative Parties “in furtherance of the principles and objectives of the Treaty.”
- Rule 2: Content of each adopted “measure” would include its text along with any tables or figures.

- Rule 3: Exclude the “*extracts*” because they were not formal “*measures*” that had been adopted by the Antarctic Treaty Consultative Parties “*in furtherance of the principles and objectives of the Treaty.*”
- Rule 4: Exclude the “introductory notes” because they were additions from the United States Department of State that had not been adopted by the Antarctic Treaty Consultative Parties “*in furtherance of the principles and objectives of the Treaty.*”

The next decision was to identify the appropriate granularity of the policy documents that would be searchable. This concept of granularity refers to the smallest conceptual units (i.e., information granules) that could be resolved directly from the overall collection. Each Antarctic Treaty Consultative Meeting produced a report with adopted “*recommendations,*” “*decisions,*” “*measures*” or “*resolutions,*” which sometimes included “*appendices,*” “*annexes*” or “*attachments.*” Periodically, the Antarctic Treaty Consultative Parties also adopted Conventions and larger policy documents that included specific “*articles*” along with “*annexes*” (Table 1). Based on these types of adopted measures, the following rules were used to create the appropriate granularity of the policy documents for the Antarctic Treaty System that could be expressed in hierarchical formats:

- Rule 5: Each “*recommendation,*” “*decision,*” “*measure*” or “*resolution*” would be treated as a complete information granule (within the context of the Antarctic Treaty Consultative Meeting and year when it was adopted).
- Rule 6: Each “*appendix,*” “*annex*” and “*attachment*” would be treated as a complete information granule (within the context of the “*recommendation,*” “*decision,*” “*measure*” or “*resolution*” within the Antarctic Treaty Consultative Meeting and year when it was adopted).
- Rule 7: Each “*article*” and “*annex*” within a Convention or larger policy document would be treated as a complete information granule (within the year that it was adopted).

Lastly, it was necessary to preserve the provenance for each of the information granules so that they could be comprehensively integrated and related objectively to each other (Figs. 2 and 3). The provenance for each information granule, which also can be referred to as a ‘concept space,’ was reflected by its unique parent-child relationship and location within the overall collection or ‘knowledge space.’ The rules to preserve and expose the provenance for each information granule were:

- Rule 8: Each information granule or ‘concept space’ would contain hierarchical information about its parent-child relationship and location within the overall collection or ‘knowledge space.’
- Rule 9: The overall collection or ‘knowledge space’ would be displayed as expandable-collapsible hierarchies that would reveal objective relationships among relevant information granules or ‘concept spaces’ for any search.

Information granules in the completed digital record of the *Antarctic Treaty Searchable Database* can be objectively integrated and comprehensively displayed in expandable-collapsible hierarchies activities (Figs. 2 and 3) that are dynamically generated for any search. These hierarchal displays expose relationships within and between the original digital elements that otherwise would be hidden in lists that are conventionally produced by search engines. The underlying technologies for the to implement the *Antarctic Treaty Searchable Database* are described below.

The general activities to create the digital record of the *Antarctic Treaty Searchable Database*, as well as similar databases of policy documents, are illustrated in Figure 4. The first step is to define the collection parameters, which includes the components of the collections as well as the resulting granularity and organization of the hierarchal displays that are dynamically generated for the integration queries. After compiling the collection elements, the next step is to implement the appropriate granularity with header tags in each granule that describe the parent-child relationships, which will be used to dynamically display the relationships among granules in the hierarchies (see the Technological section below; pages 15-18). After searching and integrating the granules, the relationships are assessed to determine whether the appropriate collections are included and the organization of the hierarchal levels meets the criteria of the user.

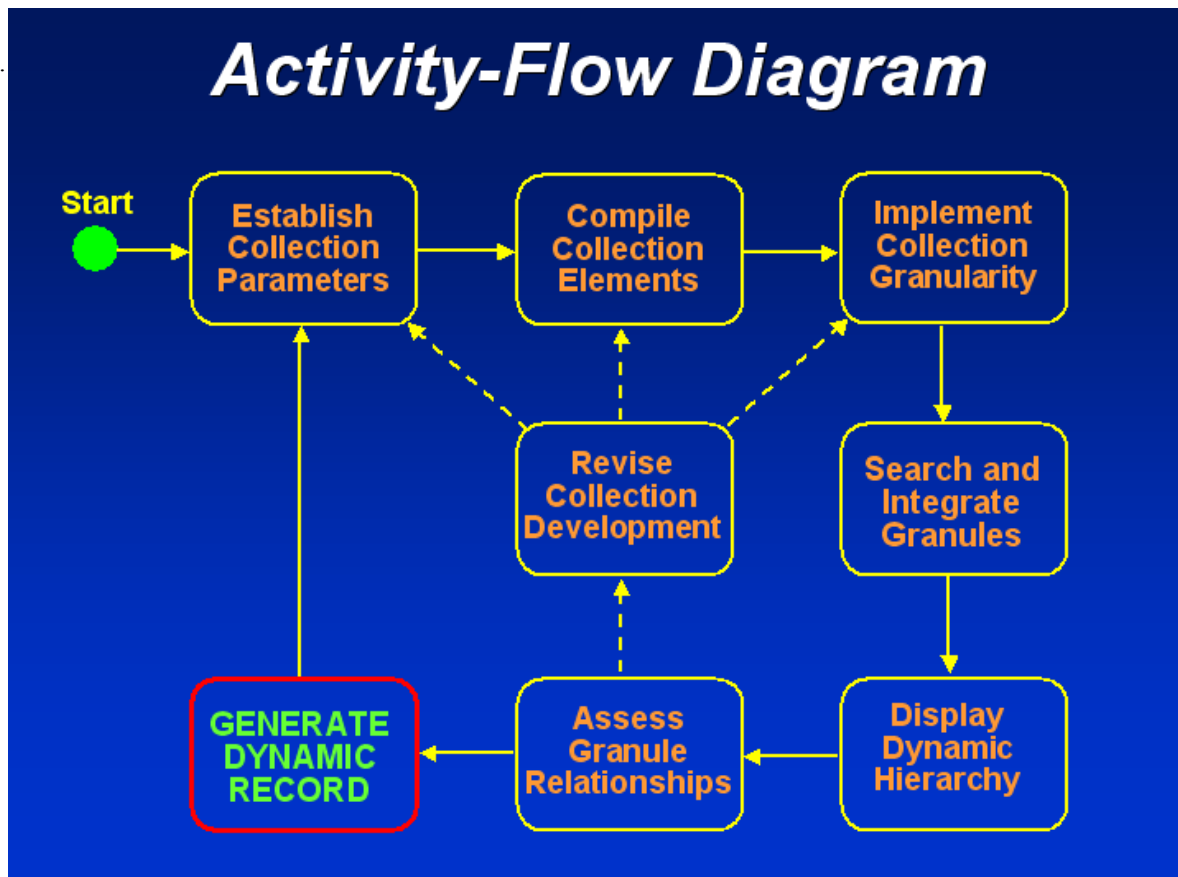


Figure 4. A generalized Activity-Flow-Diagram to illustrate the processes to create the *Antarctic Treaty Searchable Database* or other digital records with the *Digital Integration System™* (DIGIN®) from EvREsearch LTD (see pages 15-18).

The specific flow of activities and the data elements to implement the *Antarctic Treaty Searchable Database* are illustrated in Figure 5.

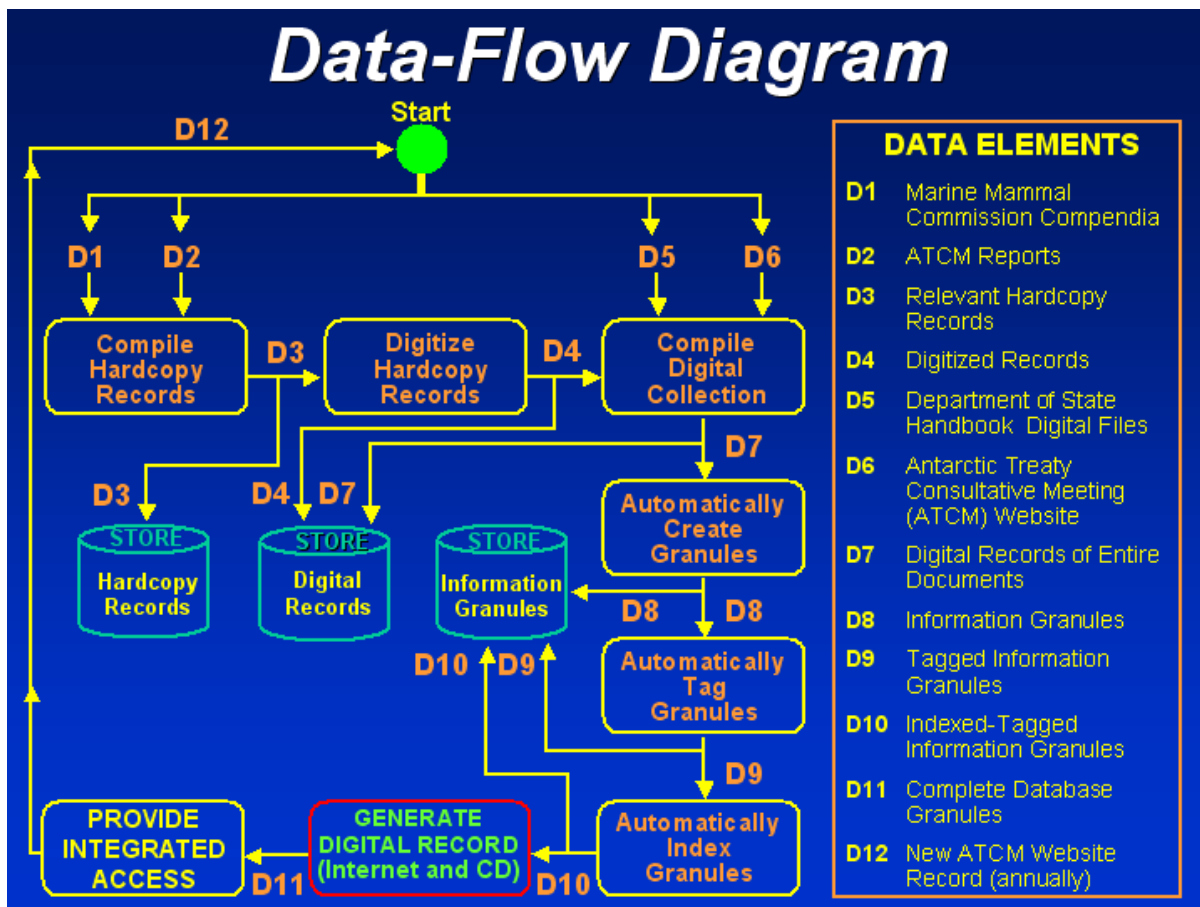


Figure 5. A Data-Flow-Diagram to illustrate the activities (Fig. 2) along with specific data elements and data stores to implement *Antarctic Treaty Searchable Database* with integrated access to the policy documents from the Antarctic Treaty System.

Documentary Context (fonds of the digital record and its internal structure)

The digital record of the *Antarctic Treaty Searchable Database*, along with all of its elements (i.e., fonds), has been continuously updated (Figs. 4 and 5) as:

- new ATCM “measures” are adopted annually by the Antarctic Treaty Consultative Parties;
- missing ATCM “measures” from the digital record are identified; and,
- missing components (e.g., tables, figures, attachments or annexes) from the measures in the digital record are identified.

Throughout, the fonds have been ingested from authentic sources (Fig. 5) and incorporated into the digital record of the *Antarctic Treaty Searchable Database* without modifying their content.

Antarctic Treaty Searchable Database also preserves the internal structures of its elements along with their provenance (as discussed above). In addition, these internal structures of the digital record are expressed by expandable-collapsible hierarchies that objectively and comprehensively integrate the information granules (Figs. 2 and 3), which are described further in the technology section below.

Rather than just accounting for the internal structure of the digital record to ensure that its element relationships are maintained, the structure provides an objective framework to implement the granularity of the digital record. Unlike hardcopy records that only can be managed in relation to their content, digital records can be managed based on both their content and structure. These two intertwined characteristics of ‘information’ are illustrated by an encrypted message that has content, but with a hidden structure that obscures any meaning. In fact, the overall internal structure of the digital record and the patterns of content in its elements provide the basis for the DIGIN[®] technologies that were used to implement the *Antarctic Treaty Searchable Database*.

Technological Context (the characteristics of the digital environment in which the record is created and maintained)

The *Digital Integration System*[™] (DIGIN[®]) from EvREsearch LTD, which was used to implement the *Antarctic Treaty Searchable Database*, involves four principal modules that can be used together or separately (Fig. 6). Each of the modules acts upon a set of expert rules that define its operation. These rule sets are optimized iteratively to objectively integrate and display units of information within and between digital resources. In addition, because DIGIN[®] is modular, it can interface with statistical, graphical, semantic web, natural language or other types of software solutions that could be treated as additional modules.

DIGIN[®] provides a general method that operates independently from any specific hardware and software. DIGIN[®] itself is written in PERL (Practical Extraction and Reporting Language), which provides a stable cross-platform programming language that can read and write binary files as well as process very large files. In addition, ASCII (American Standard Code for Information Interchange) format is used because it is a world-wide standard for representing for Latin text, numbers, punctuations and symbols. Consequently, because DIGIN[®] is an interoperable method that is independent of any specific file format, it can be utilized into the future in a persistent manner.

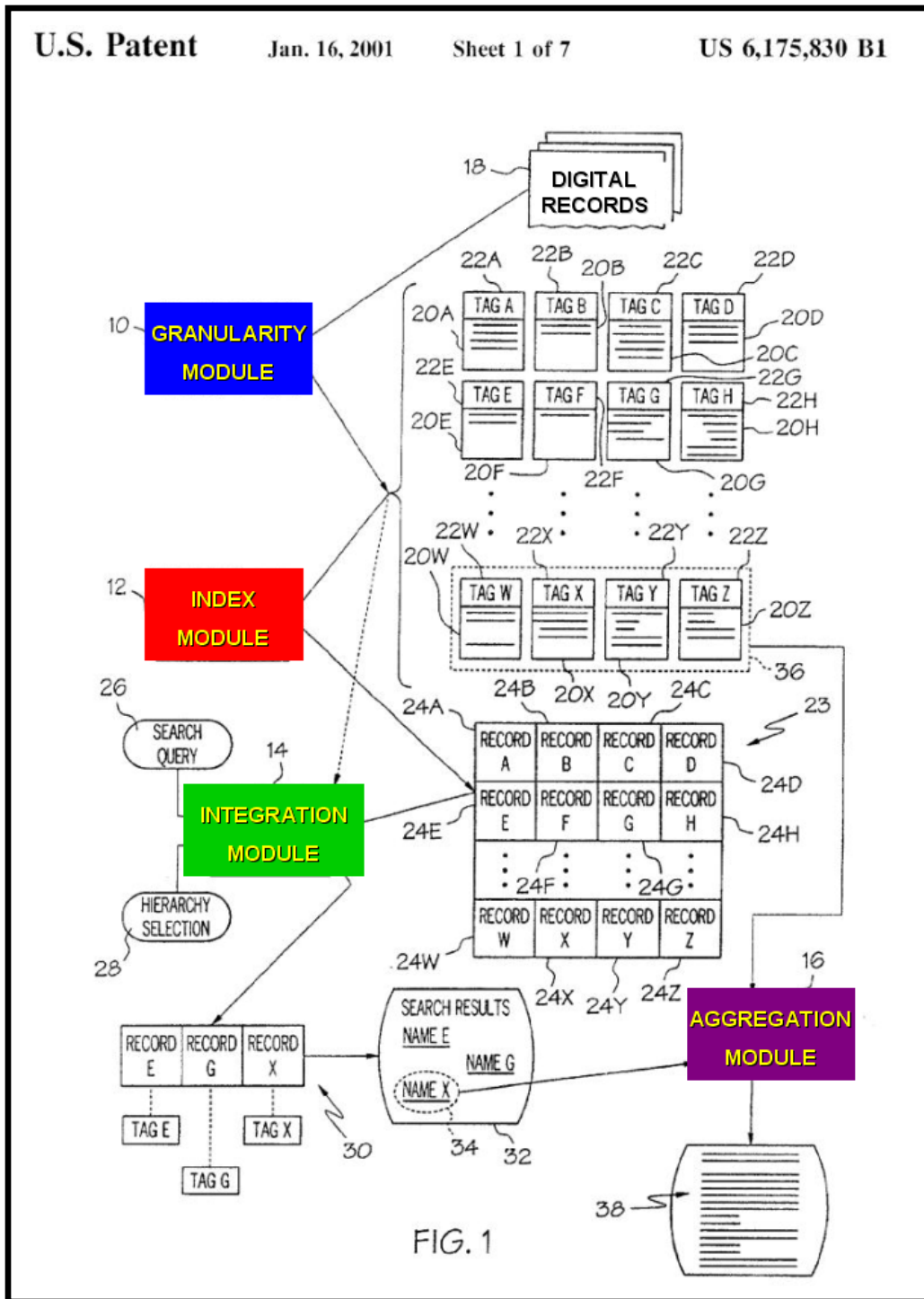


Figure 6. The *Digital Integration System™* (DIGIN®) from EvREsearch LTD is based on the “*Information Management, Retrieval and Display Systems and Associated Methods*” (United States Patent Nos. 6,175,830 and 6,484,166; New Zealand Patent No. 515007; Australian Patent No. 770087; Pending in Canada, China, European Union, India, Japan and Mexico) that are illustrated in the schematic. This modular system, as described in the text, utilizes the inherent structural patterns in digital records to automatically generate information granules that can be dynamically integrated, displayed and aggregated from expandable-collapsible hierarchies based on user-defined criteria.

The four DIGIN[®] modules are:

- *GRANULARITY MODULE*: automatically creates the information granules based on the inherent structural patterns or boundaries of the information units within the digital record groups, series and entities. A unique categorical tag is assigned to each granule based upon an analysis (defined by the set of expert system rules) of its provenance, parent-child location and contents. The categorical tags contain information to generate expandable-collapsible hierarchies.
- *INDEX MODULE*: automatically generates a database with the address (referenced within each categorical tag), content strings (words, numbers or other symbols) and their frequencies within each information granule.
- *INTEGRATION MODULE*: comprehensively searches through the index for the information granules with terms or strings that match the user-defined search queries in textual, numeric or other symbolic forms. By applying the categorical tags and/or weight of the matching search strings, the relevant information granules are then integrated and displayed objectively in expandable-collapsible hierarchies.
- *AGGREGATION MODULE*: combines relevant information granules based on their hierarchal relationships and user-defined criteria. The granules can be aggregated to reconstruct contiguous portions of original information resources or to create new information resources.

Unlike subjective content descriptions in metadata or controlled vocabularies, DIGIN[®] comprehensively searches both the contents of the granules and their categorical tags to objectively identify those granules that match the search queries. Moreover, there is no ambiguity about a null search result with these analyses. Consequently, the objective hierarchal relationships that are dynamically identified from a DIGIN[®] integration (e.g., Figs. 2 and 3) can be quantified (Fig. 7).

Moreover, DIGIN[®] is interoperable with metadata, mark-up and databases, which each have standardized patterns that can be easily described in rule sets to initiate the GRANULARITY MODULE. This automated technology neither relies on nor requires these conventional information-management technologies, however, to discover relationships among digital resources.

The capacity to comprehensively search the granules and develop relational displays (e.g., Figs. 2, 3 and 7) provides an objective framework to interpret trends that ultimately facilitate decision-making. With the *Antarctic Treaty Searchable Database*, this functionality has practical value for academic research and education as well as industry, government and other stakeholders with interests regarding the Antarctic Treaty region “*south of 60° south latitude.*”

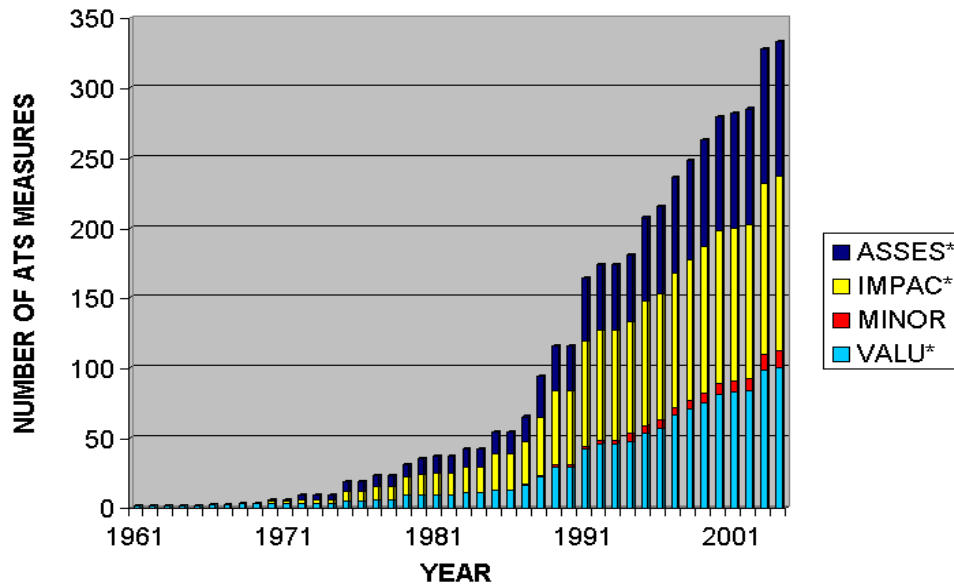


Figure 7. Cumulative profile of Antarctic Treaty “measures” with search terms (* is the wildcard search character) that were comprehensively integrated over time with the *Antarctic Treaty Searchable Database* (<http://aspire.tierit.com/>), as expressed by hierarchal displays (e.g., Figs. 2 and 3).

The digital record of the *Antarctic Treaty Searchable Database* has expanded from 608 to 740 granules between March 1999 and September 2004. However, throughout this period, the *Antarctic Treaty Searchable Database* has been maintained with the same uniform resource locator on the Internet (<http://aspire.tierit.com/>). Each of the annual editions of the *Antarctic Treaty Searchable Database* is preserved on webCDservers that contain fully executable copies of the Web sites.

D. Addressing the 23 Core Research Questions

1. What activities of the creator have you investigated?

As the creator, the following activities associated with the digital record of the *Antarctic Treaty Searchable Database* (<http://aspire.tierit.com/>) have been investigated:

- a. establishing parameters of the digital record;
- b. compiling elements of the digital record;
- c. generating granularity of the digital record;
- d. updating the digital record;
- e. applying the patented *Digital Integration System*TM (DIGIN[®]);
- f. searching and displaying granules in the digital record;
- g. integrating the granules in the digital record;
- h. quantifying relationships among granules in the digital record;
- i. utilizing the digital record in educational settings;
- j. disseminating the digital record;
- k. assessing applications of the digital record; and
- l. comparing technology solutions to access policy documents that have been adopted by the Antarctic Treaty Consultative Parties.

These activities are elaborated in the above sections on the Description of Digital Entity Creation.

2. Which of these activities generate the digital entities that are the objects of your case study?

Activities a-e (from question 1 above) are involved in the generation of the digital entities in the *Antarctic Treaty Searchable Database*. These digital entities refer to components or information granules in the digital record, which is considered to be the *Antarctic Treaty Searchable Database* itself in this case (see question 18). These digital entities are self-contained concept spaces that can be objectively defined and extracted in relation to the structural organization of their parent information resources. In addition, these digital record entities represent higher levels of granularity than the digital record series and digital record groups that characterize the collection of information resources in the overall knowledge space.

After establishing the collection parameters (Figs. 4 and 5), based on the rule sets above (pages 11-12), the elements were compiled and then automatically broken into separate information granules using the *Digital Integration System*TM (DIGIN[®]). In particular, the Granularity Module (Fig. 6) was the principal tool for generating the digital entities that are the objects in *Antarctic Treaty Searchable Database Case Study*.

3. For what purpose(s) are the digital entities you have examined created?

Initially, in 1999, the digital record of the *Antarctic Treaty Searchable Database* was created for educational purposes (Berkman 2002). The digital record has been updated and maintained subsequently for both educational and non-educational activities. The non-educational users include international government organizations, national government agencies, non-governmental organizations and industry. The primary objective currently is to implement the *Antarctic Treaty Searchable Database* in a manner that:

- (a) contributes to the operation of the Antarctic Treaty Secretariat; and
- (b) could be sustained by the Antarctic Treaty Secretariat into the future for the broad benefit of educational and non-educational audiences.

The principal reason for creating the digital entities is to enhance the potential to discover and quantify knowledge about Antarctic policies in an objective manner. By increasing the level of granularity beyond the parent documents, there is greater capacity to integrate concepts and discover meaningful relationships. For example, two objects have three possible relationships whereas four objects have fifteen possible relationships and eight objects have 255 possible relationships (Fig. 8).

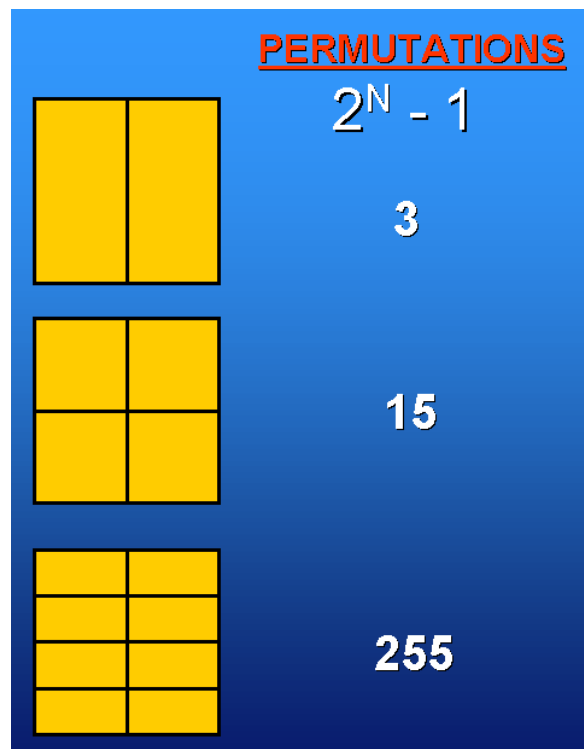


Figure 8. Illustration of the exponentially increasing number of possible permutations or relationships among 'N' objects.

In the *Antarctic Treaty Searchable Database* there are nearly 750 digital entities, as opposed to the twelve ‘pdf’ files on the United States Department of State Web site that contain the Antarctic Treaty documents from 1959 through 2002 (Box 1). Because the digital record entities (i.e., the twelve ‘pdf’ files) on the Department of State Web site and CD are both fewer and subjectively organized, there also is limited capacity to discover relationships without manually extracting information from each of the ‘pdf’ files separately. Moreover, the twelve files on the Department of State Web site are out of date. In contrast, because the granules are objectively generated and can be comprehensively integrated (Figs. 2 and 3), the *Antarctic Treaty Searchable Database* provides an automated framework to turn qualitative information into quantitative results (Fig. 7) for subsequent interpretation and decision making by diverse audiences.

4. What form do these digital entities take? (e.g., e-mail, CAD, database)

The digital record entities are individual files with text in ASCII format and inserted figures (saved as “jpg” files). The objectives of utilizing ASCII files are to normalize the file formats and to utilize an open-standard file format that will persist through time.

The digital entities are a collection of information granules that can be dynamically, comprehensively and objectively integrated based on their inherent structural relationships (see pages 15-18) without conventional metadata, markup or databases. Although the term database is used in the name *Antarctic Treaty Searchable Database*, the digital collection integrates information granules without cells that intersect rows and columns that conventionally are associated with the operation of a “database.”

4a. What are the key formal elements, attributes, and behaviour (if any) of the digital entities?

As specified in the Data-Flow-Diagram (Fig. 5), the key formal elements of the digital entities are “*measures in furtherance of the principals and objectives of the [Antarctic] Treaty*” that have been adopted by the Antarctic Treaty Consultative Parties between 1959 and the present. The key formal elements in the *Antarctic Treaty Searchable Database* originated in both hardcopy and digital formats.

The key formal elements of the digital record entities include the adopted measures that are included in the final reports from each of the Antarctic Treaty Consultative Meetings from 1961 through the 2004. In addition, the key formal elements include the individual articles within the Antarctic Treaty and its Conventions (Table 1). The digital record entities or granules represent the final level in a hierarchy underlying the record series of adopted measures from each year within the hierarchy where the record group is the Antarctic Treaty System (see question 2).

4b. What are the digital components of which they consist and their specifications?

The specifications for the digital components of the *Antarctic Treaty Searchable Database* are defined by Rules 1-7 above (pages 11-12). These digital components originated in digital form or were scanned from hardcopy into digital forms. The scanned digital components largely involved figures that were saved as ‘jpg’ files and then inserted into the appropriate digital entities. The specific digital entities, which represent a higher level of granularity than their overlying digital record series and digital record group (see question 4a), include:

- articles;
- measures;
- decisions;
- resolutions;
- annexes; and
- appendices.

4c. What is the relationship between the intellectual aspects and the technical components?

The intellectual aspects (i.e., the content and context) of the *Antarctic Treaty Searchable Database* is public domain. The technical components to facilitate the functionality of the *Antarctic Treaty Searchable Database*, which has been freely accessible over the Internet, are illustrated in Figure 6 and described above (pages 15-18).

4d. How are the digital entities identified (e.g., is there a [persistent] unique identifier)?

Each of the information granules or digital record entities in the current *Antarctic Treaty Searchable Database* contains its unique provenance information in a categorical header tag as well as its title. Unlike metadata, which are stored in repositories separately from the digital entities, the unique identifiers are part of each granule in the *Antarctic Treaty Searchable Database*. Consequently, with the categorical header tags, there is never a risk for decoupling the unique identifiers and the digital entities. The categorical header tag operates with the Integration Module to dynamically create the expandable-collapsible hierarchies. An example of the header tag for a digital record entity from the *Antarctic Treaty Searchable Database* is shown below (Box 2).

4e. In the organization of the digital entities, what kind of aggregation levels exist, if any?

The aggregation levels among digital entities in the *Antarctic Treaty Searchable Database* are based on the inherent parent-child relationships within the policy documents. In general, the aggregation levels or hierarchy levels (e.g. Figs. 2 and 3) reflect the granularity of a digital collection (Box 3).

BOX 2
**AN EXAMPLE OF THE UNIQUE IDENTIFIERS ASSOCIATED WITH EACH
DIGITAL RECORD ENTITY IN THE *ANTARCTIC TREATY SEARCHABLE*
DATABASE**
(<http://aspire.tierit.com/>)

Header Tag

```
</head> <meta name="Dewey" content="2004%%====%%Twentyseventh  

Consultative Meeting%%====%%Decisions%%====%%%%====%%Decision  

XXVII-2 Financial Considerations for the Secretariat of the Antarctic Treaty">
```

Title in the Digital record Entity

ATCM XXVII (2004)
Decision XXVII-2
[*Financial Considerations for the Secretariat of the Antarctic Treaty*]

BOX 3
**GENERALIZED HIERARCHAL FRAMEWORK OF AN ARCHIVED DIGITAL
RECORD**

Digital record Group > Digital record Series > Digital record Entities

This collection granularity is represented specifically for the *Antarctic Treaty Searchable Database* by:

- *Antarctic Treaty Searchable Database* > Year > Major Document or Antarctic Treaty Consultative Meeting > “*measures*”

Dynamic aggregation of digital record entities with DIGIN[®] facilitates the discovery of relationships within and between the digital record series.

4f. What determines the way in which the digital entities are organized?

Organization of the digital entities (see question 4e) is objective and defined by the inherent structure, patterns and organization of their parent digital record series (i.e., there is no subjective interpretation about organization of the digital entities). The unique parent-child relationships for each digital record entity, which are preserved through its header tag and title (Box 2), are expressed through the expandable-collapsible hierarchies that are dynamically generated in response to a search query (e.g., Figs. 2 and 3).

With the DIGIN[®] system, specific implementation of the hierarchies is based on objective relationships among digital record entities within and between their digital record series. Objective levels in these hierarchies can be arranged according to user-defined criteria, which in the case of the Antarctic Treaty Searchable Database, are defined over time by Rules 8 and 9 (page 12).

5. How are those digital entities created?

The digital record entities are created from digital record series (see questions 2 and 4e) based on processes that are described in the Activity-Flow Diagram (Fig. 4). Based on the specific collection elements in the *Antarctic Treaty Searchable Database* (Fig. 5), the granularity of the digital record series was automatically increased with the DIGIN[®] system (Fig. 6) to create the digital record entities based on Rule Sets 1-7.

5a. What is the nature of the system(s) with which they are created? (e.g., functionality, software, hardware, peripherals, etc.)

The technological procedures to create and integrate the digital record entities are illustrated in Figure 6 and discussed above (pages 15-18). In particular, the Granularity Module of DIGIN[®] is used to objectively create the digital record entities (see questions 2 and 4a-f). The premise of these operations is that all information has content, context and structure. In a hardcopy world, information is managed based on its content (libraries) and context (archives). In a digital world, it also is possible to automatically manage information based on its structure. Based on user-defined criteria, this automated technology utilizes the structure of the digital information first to objectively manage the information so that the content can be comprehensively integrated and automatically expressed by hierarchal relationships subsequently.

DIGIN[®] is interoperable with databases, markup and metadata that all have well-defined patterns and structures to create rule sets for the operation of the Granularity Module (Fig. 6). However, DIGIN[®] also can operate with the pattern and structure in the original authentic digital records databases without ever applying markup, metadata and databases. Consequently, DIGIN[®] can be used interchangeably with both “structured” and “unstructured” information to facilitate the automated integration of digital record entities at user-defined levels of granularity.

Metadata is a ubiquitous tool for digital information management, yet it is not necessary for the operation of the DIGIN[®] system. First and most-importantly, metadata is an inappropriate tool for user-defined levels of granularity beyond the digital record group and the digital record series. Practically, metadata has fixed fields that must be populated independent of the size of the information granule. This means that as the granularity is increased (i.e., each granule becomes smaller), the volume of metadata increases relative to the size of each granule (Fig. 9).

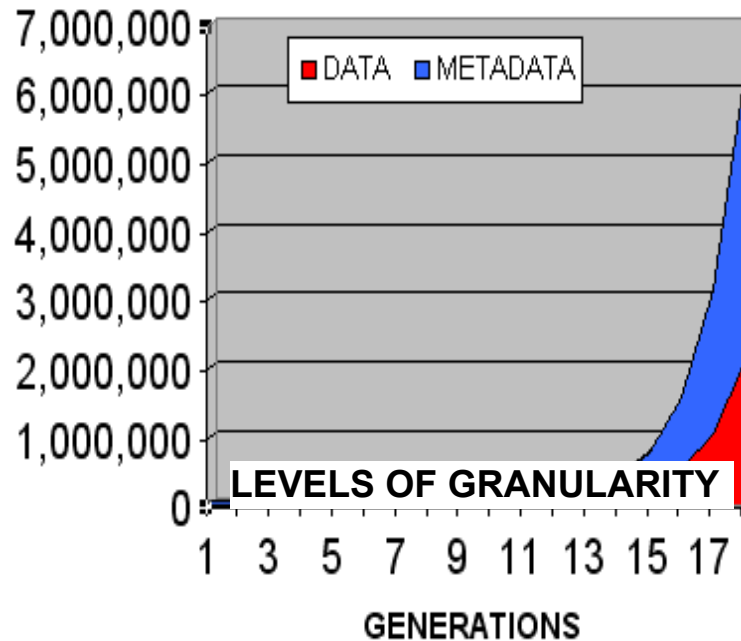


Figure 9. A model of the exponentially increasing volume of metadata by simply doubling the number of information granules (data), each of which is half its original size, while the volume of each metadata record is fixed independent of granule size.

Moreover, the fields in the metadata tags require input that involves subjective decisions, as with controlled vocabularies. These features of metadata, which are exacerbated at higher levels of granularity, influence the:

- storage and processing efficiencies;
- connections between metadata and their corresponding digital record entities; and
- objectivity of analytical results that can be quantified.

Aside from the practical limitations of metadata, the automated granularity that can be achieved with the DIGIN[®] system (Berkman and Morgan 2003) facilitates the generation of digital record entities.

By automating the generation and integration of digital record entities at user-defined levels of granularity, based on the inherent structure and organization of their parent digital record series and digital record groups (see questions 2 and 4e), DIGIN[®] provides both the convenience of search engines and the control of databases, metadata and markup.

- Search engines are convenient and very powerful tools for accessing digital information, but they provide limited control when thousands of ‘hits’ are returned and only the top few are utilized. Moreover, search engines conventionally generate lists that hide relationships within and between digital

record groups and digital record series. In addition, ranking in these lists is based on algorithms that are determined by the programmers.

- On the other end of the spectrum, databases, metadata and markup provide great control to manage digital information, but they are not convenient. Moreover, these digital management tools are subjectively designed around the content of the digital resources.

This convenience and control of the DIGIN[®] system (pages 15-18) is best represented by the questions that can be personally addressed to define its automated operations and applications (Box 4).

BOX 4

User-Defined Criteria for Automated Operation of the DIGIN[®] System (Fig. 6): Convenience and Control for Digital record Management

User-Defined Criteria

- **What are the components of the digital collection?**
- **What is the appropriate level of granularity to discover meaningful relationships in the digital collection?**
- **What are the search queries that will be imposed on the digital collection?**
- **How will digital record entities from different levels in the information hierarchies be identified, organized and displayed?**
- **What are the appropriate statistical and graphical analyses to impose on the quantitative results from a DIGIN[®] integration?**

- 5b. Does the system manage the complete range of digital entities created in the identified activity or activities for the organization (or part of it) in which they operate?**

Yes. Operation of the *Antarctic Treaty Searchable Database* involves the complete collection of digital record entities and their contents that are managed with the Index Module (Fig. 6). Consequently, queries with the *Antarctic Treaty Searchable Database* dynamically generate hierarchal displays that only reflect the relationships among digital record entities that match the search criteria (Figs. 2 and 3). Because these integration results are comprehensive, they can be quantified (Fig. 7).

6. From what precise process(es) or procedure(s), or part thereof, do the digital entities result?

The precise processes and procedures to generate the digital record entities with the DIGIN[®] system (Fig. 6) are describe above (See questions 2, 4e and 5). In addition, the activities and elements to generate the digital record entities in the *Antarctic Treaty Searchable Database* are described in the Activity-Flow Diagram (Fig. 4) and Data-Flow-Diagram (Fig. 5), respectively.

7. To what other digital or non-digital entities are they connected in either a conceptual or a technical way? Is such connection documented or captured?

The information granules in the *Antarctic Treaty Searchable Database* derive from the Antarctic Treaty and final reports from the Antarctic Treaty Consultative Meetings (as described in Fig. 5) as well as the other Antarctic policy documents that are described in Table 1. Origin of these information granules came from both digital and non-digital formats.

In addition, the entire scope of the Antarctic Treaty Searchable Database is within the context of international law and other legal regimes. For example, the scope of the Antarctic Treaty is defined within Article VI, which refers to the “*area south of 60° south latitude*” that also is “*under international law with regard to the high seas within that area.*” Consequently, in this example, the digital entities in the *Antarctic Treaty Searchable Database* also are connected to the 1958 *Convention on the High Seas* as well as the 1984 *United Nations Convention on the Law of the Sea* (<http://lawofthesea.nvi.net>) and other marine-related regimes. The digital entities are further connected to other international regimes that are specified, which include the:

- *1966 International Convention on Load Lines;*
- *1972 Convention on the International Regulations for Preventing Collisions at Sea;*
- *1972 Convention for the Prevention of Marine Pollution by the Dumping of Wastes and other Matter;*
- *1973 International Convention for the Prevention of Pollution from Ships;*
- *1974 International Convention for the Safety of Life at Sea;*
- *1978 International Convention on Standards of Training, Certification and Watchkeeping for Seafarers with Annex; and,*
- *1989 Basel Convention on the Control of Transboundary Movements of Hazardous Wastes and their Disposal.*

8. What are the documentary and technological processes or procedures that the creator follows to identify, retrieve, and access the digital entities?

The documentary process to identify, retrieve and access the digital entities in the *Antarctic Treaty Searchable Database* are described by the Data-Flow-Diagram (Fig. 4.). The technological processes to identify, retrieve and access the digital entities are illustrated in Figure 6 and described above (pages 15-18).

9. Are those processes and procedures documented? How? In what form?

Yes. The documentary processes to implement the *Antarctic Treaty Searchable Database* are describe herein. The technological processes are described in the:

- patented “*Information Management, Retrieval and Display Systems and Associated Methods*” ((United States Patent Nos. 6,175,830 and 6,484,166; New Zealand Patent No. 515007; Australian Patent No. 770087; Pending in Canada, China, European Union, India, Japan and Mexico);
- *report on “Automated Granularity of Authentic Digital Records in a Persistent Archive,”* which was prepared for the National Archives and Records Administration, is available through San Diego Supercomputer Center Web site (<http://www.sdsc.edu/NARA/Publications.html>); and
- abstracts of papers that presented in 2004 to formal committees through the International Council of Science and National Science Digital Library program coordinated by the National Science Foundation (see Appendix 1: Abstracts).

10. What measures does the creator take to ensure the quality, reliability and authenticity of the digital entities and their documentation?

To ensure reliability, the content and context of the *Antarctic Treaty Searchable Database* is preserved and unmodified from its original sources, which are described in Figure 5. Ongoing correspondence with members from the Antarctic Treaty Consultative Parties is outlined in Table 2. Reliability of the *Antarctic Treaty Searchable Database* and its practical value are best reflected by the diverse stakeholders around the world who have created Web site linkages to this digital collection of integrated policy documents (Box 5).

For the webCDserver, documentation about the *Antarctic Treaty Searchable Database* is described in the “*End Users License Agreement*” that must be agreed it can be accessed. With the Web site, to ensure that the accuracy and authenticity of this digital record are clearly understood, the following disclaimer is used:

This searchable database is being developed to enhance public access to the Antarctic Treaty documents and any errors or omissions are included from the original versions from the United States Department of State and National

*Science Foundation. There is no warranty, expressed or implied, as to the accuracy or completeness of the furnished data or the resulting searchable database displays.*⁵

All versions of the *Antarctic Treaty Searchable Database* also have a “Read-Me” file that describes its implementation.

11. Does the creator think that the authenticity of his digital entities is assured, and if so, why?

No. Authenticity, in the case of public-domain policy documents, can only be assured by the government agencies that issue the records. With measures from the Antarctic Treaty System, which are international policy documents, the authentic versions are those transmitted by formal diplomatic channels in hardcopy formats by accepted government authorities, which include the Department of State for the United States as depository government and the Antarctic Treaty Secretariat. Nonetheless, the *Antarctic Treaty Searchable Database* is accurate and reliable as reflected by its applications for diverse stakeholders around the world (Box 5). Warranties and disclaimers about the accuracy and authenticity of the *Antarctic Treaty Searchable Database* are described in question 10.

12. How does the creator use the digital entities under examination?

The *Antarctic Treaty Searchable Database* was created in 1999 because students in the interdisciplinary *Antarctic Marine Ecology and Policy* capstone course, which had been taught by Paul Berkman since 1982, were unable to effectively locate and digest documents from the 1000-page *Antarctic Treaty Handbook* published by the United States Department of State. Subsequently, the searchable database has provided an open-ended inquiry framework for students as well as diplomats, managers of national Antarctic programs, tourists and other individuals (e.g., Box 4) to interpret the international policies surrounding human activities in Antarctica.

The webCDserver with the *Antarctic Treaty Searchable Database* also was an essential innovation for the textbook on *Science into Policy: Global Lessons from Antarctica* that was published by Academic Press (Berkman, 2002). The Web site and webCDserver with the *Antarctic Treaty Searchable Database* provided the first framework for integrating policy documents from the Antarctic Treaty System. The *Antarctic Treaty Searchable Database* also remains as the most comprehensive source for integrating the “*measures*” that have been adopted by the Antarctic Treaty Consultative Parties between 1959 and the present.

⁵ <http://aspire.tierit.com/enter.htm>.

Table 2. Chronology of international communications regarding the Antarctic Treaty Searchable Database following its introduction by the United States Department of State at the 23th Antarctic Treaty Consultative Meeting in Peru in May 1999⁶

Date	Activity	Nation
February 2001	E-Mail Message from National Delegate	Australia
January 2002	E-Mail Message from National Delegate	New Zealand
January 2002	Meeting with National Delegate in Christchurch	New Zealand
February 2002	Meeting with National Delegate in Hobart	Australia
February 2002	E-Mail Message from National Delegate	Australia
April 2002	E-Mail Message from National Delegate	Australia
August 2002	E-Mail Message from National Delegate	The Netherlands
August 2002	E-Mail Message from National Delegate	Australia
September 2002	E-Mail Message from National Delegate	Spain
September 2002	E-Mail Message from National Delegate	Australia
October 2002	E-Mail Message from National Delegate	The Netherlands
April 2003	E-Mail Message from National Delegate	Spain
May 2003	International Mailing of webCDservers 3rd Edition	
May 2003	E-Mail Message from National Delegate	Japan
May 2003	E-Mail Message from National Delegate	South Korea
May 2003	E-Mail Message from National Delegate	Germany
June 2003	E-Mail Message from National Delegate	India
June 2003	E-Mail Message from National Delegate	New Zealand
August 2003	E-Mail Message from National Delegate	Bulgaria
September 2003	E-Mail Messages from National Delegate	The Netherlands
October 2003	Meeting with Dutch National Delegate in Washington D.C.	
October 2003	Meeting with Argentine National Delegate in Buenos Aires	
January 2004	International Mailing of webCDservers 4th Edition	
February 2004	E-Mail Message from National Delegate	Argentina
February 2004	E-Mail Message from National Delegate	Japan
February 2004	E-Mail Message from National Delegate	New Zealand
February 2004	E-Mail Message from National Delegate	Brazil
February 2004	E-Mail Message from National Delegate	Germany
February 2004	E-Mail Message from National Delegate	France
June 2004	E-Mail Message from National Delegate	New Zealand
July 2004	E-Mail Delivery of 26 th ATCM Documents	South Africa
July 2004	Meeting with Antarctic Treaty Executive Secretary in Bremen	
July 2004	Distribution of webCDservers 5th Edition in Bremen	
December 2004	Expression of interest from the Antarctic Treaty Secretariat to utilize the <i>Antarctic Treaty Searchable Database</i> technology	

⁶ Ongoing communications with United States Department of State are not included.

BOX 5**REPRESENTATIVE WEBSITE LINKS TO THE
ANTARCTIC TREATY SEARCHABLE DATABASE
(<http://webhost.nvi.net/aspire>)****INTERNATIONAL GOVERNMENT INSTITUTIONS**

Antarctic Treaty Consultative Meeting XXIV (St. Petersburg, Russia)

<http://www.24atcm.mid.ru/>

Antarctic Treaty Consultative Meeting XXV (Warsaw, Poland)

<http://www.25atcm.gov.pl>

Antarctic Treaty Secretariat

<http://www.ats.org.ar/links.htm>

NATIONAL GOVERNMENT AGENCIES

Australian Antarctic Division

<http://www-aadc.antdiv.gov.au/>

Canadian Department of Foreign Affairs and International Trade

<http://www.dfait-maeci.gc.ca/circumpolar/antarctica-en.asp>

Library of Congress

www.loc.gov/rr/international/frd/government_law.htm

National Academy of Sciences

http://www7.nationalacademies.org/prb/Arctic_and_Antarctic_Links.html

NON-GOVERNMENTAL ORGANIZATIONS

Antarctic Southern Ocean Coalition

<http://www.asoc.org/links.htm>

International Polar Heritage Committee

<http://www.polarheritage.com/index.cfm/RefmatOtherPolar>

Joint Committee on Antarctic Data Management

<http://www.jcadm.scar.org/links1.html>

Scientific Committee on Antarctic Research

<http://www.scar.org/information/links/>

CORPORATIONS

American Society of International Law

<http://users.erols.com/jackbobo/>

Association of American Geographers

<http://www.aag.org/geotask/>

International Association of Antarctica Tour Operators

<http://www.iaato.org/>

UNIVERSITIES

George Washington University Law School

<http://www.law.gwu.edu/burns/research/intl/env.htm>

Katholieke Universiteit Leuven

<http://www.kuleuven.ac.be/iir/linkse.htm>

Oxford University

<http://www.oup.uk/pdf/bt/cassese/cases/part1/ch03/614.pdf>

Texas A&M

http://www.gerg.tamu.edu/antarctica/menu_links/links.htm

University of California, Santa Barbara

<http://fiesta.bren.ucsb.edu/~gsd/links/links.php?nav=nonprofit>

An application with the *Antarctic Treaty Searchable Database* to objectively and comprehensively integrate information can be illustrated in relation to environmental protection. With the *1991 Protocol on Environmental Protection to the Antarctic Treaty* (Table 1), “*minor or transitory impacts*” became the central concepts for assessing human activities in the Antarctic Treaty area. Minor is a subjective term that reflects values, whereas transitory is an objective term that reflects rates. Consequently, the following terms can be used to interpret the development of environmental protection strategies in the Antarctic Treaty System:

- “minor”
- “impact”
- “assess”
- “value”

Figures 2 and 3 provide examples of expandable-collapsible hierarchies. Since the displays comprehensively identify the occurrences of specific search terms in the policy documents over time, these results can be quantified. Figure 7 illustrates the graphical display of the cumulative frequencies of “minor” and other search terms over time. To further interpret these results, it can be hypothesized that environmental protection concepts have become progressively more integrated over time as the Antarctic Treaty Consultative Parties negotiated environmental protection strategies. This hypothesis was tested by quantifying the relationships among the search terms over time (Fig. 10).

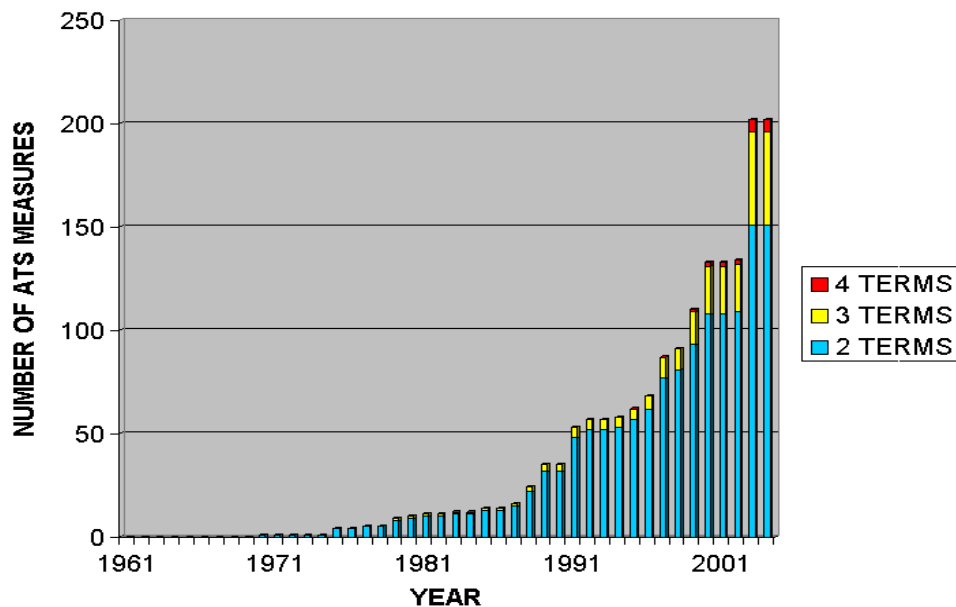


Figure 10. Cumulative profile of relationships among Antarctic Treaty “*measures*” with 2, 3 or 4 of the search terms from Figure 6 that were comprehensively integrated over time with the *Antarctic Treaty Searchable Database*, as expressed in hierarchal displays (e.g., Figs. 2 and 3).

Figure 10 shows that the Antarctic Treaty “measures” first integrated 2 then 3 and finally all 4 of the search terms. Also, the frequencies of “measures” with 2, 3 and 4 terms increased over time. These results support the idea (i.e., they do not falsify the hypothesis) that the underlying concepts of Antarctic environmental protection became progressively integrated over time. This illustration demonstrates how the DIGIN[®] technologies can be used to integrate and objectively turn qualitative information into quantitative data. This capacity to objectively integrate information and quantify trends has significant applications for making decisions based on disparate information resources, as is being identified in the topical areas of water conflict resolution (Yoffe et al. 2004, Song and Whittington 2004).

13. How are changes to the digital entities made and recorded?

After they are implemented, content of the digital record entities is preserved and unmodified over time. Content and hierarchal information about the relative location of each digital record entity in the *Antarctic Treaty Searchable Database* are managed and recorded with the Index Module of the DIGIN[®] system (see question 5b).

14. Do external users have access to the digital entities in question? If so, how, and what kind of uses do they make of the entities?

Yes. Users have open access to the *Antarctic Treaty Searchable Database* through the Web site as well as the distributed webCDservers. Applications of the *Antarctic Treaty Searchable Database* (as suggested by Web site links in Box 5) include:

- decision-making by government agencies at international and national levels;
- advocacy for policy changes;
- responses to existing policies for commercial purposes; and,
- consideration of policy concepts and relationships for education activities.

15. Are there specific job competencies (or responsibilities) with respect to the creation, maintenance, and/or use of the digital entities? If yes, what are they?

Responsibilities for managing the *Antarctic Treaty Searchable Database* are to:

- update it annually as “measures” from the Antarctic Treaty Consultative Parties are approved;
- preserve the accurate contents in each of the information granules without changes;
- maintain the *Antarctic Treaty Searchable Database* on a publicly- accessible Web site;
- communicate with stakeholders (e.g., Table 2) who are utilizing the *Antarctic Treaty Searchable Database*;
- distribute updated editions of the webCDserver; and
- respond to the needs and interests of the Antarctic Treaty Secretariat as well as other government programs involved with the Antarctic Treaty System.

These responsibilities have been assumed without formal mandate by Paul Berkman and EvREsearch LTD, who conceived and have been continuously implementing the *Antarctic Treaty Searchable Database* (Berkman 2002) to integrate adopted “*measures*” from the Antarctic Treaty System “*in the interest of all mankind,*” as promoted in the *Preamble* of the *1959 Antarctic Treaty*.

It is hoped that the Antarctic Treaty Secretariat will take on these responsibilities to sustain the *Antarctic Treaty Searchable Database* in a manner that best benefits the diverse community of users around the world who have been applying this integrated database (Box 5).

16. Are the access rights (to objects and/or systems) connected to the job competence of the responsible person? If yes, what are they?

The *Antarctic Treaty Searchable Database* is open access. Modifications of the *Antarctic Treaty Searchable Database* and associated applications of the DIGIN[®] technologies are permitted with approval by EvREsearch LTD (see question 15).

17. Among its digital entities, which ones does the creator consider to be records and why?

The *Antarctic Treaty Searchable Database* is itself a digital record. Moreover, the *Antarctic Treaty Searchable Database* includes the digital record entities that have been created by the DIGIN[®] system with the Granularity Module (see question 5a). As stated in question 2, the *Antarctic Treaty Searchable Database* is analogous a digital record group with underlying digital record series and digital record entities that can be expressed in hierarchies (Box 3). Consequently, the *Antarctic Treaty Searchable Database* mirrors the components of an archive. These assertions are based on the characteristics of a digital record as defined in the Glossary:

A document becomes a record when it is placed in an organized filing system for use as evidence or information. It becomes archival when transferred to a repository for preservation and research use.

18. Does the creator keep the digital entities that are currently being examined? That is, are these digital entities part of a record keeping system? If so, what are its features?

Yes. Along with the *Antarctic Treaty Searchable Database* and all of its digital record entities, which it hosts, EvREsearch LTD maintains the data stores that are illustrated in the Data-Flow-Diagram (Fig. 5).

The *Antarctic Treaty Searchable Database* operates without formal Electronic Recordkeeping (ERK) or Electronic Records Management (ERM) systems, as described by the National Archives and Records Administration.⁷ Nonetheless, the *Antarctic Treaty*

⁷ See http://www.archives.gov/records_management/policy_and_guidance/examples_system_functions.html.

Searchable Database does meet the general requirements of ERK or ERM systems by allowing users to print and view all system management and control information along with the records themselves. In addition, the *Antarctic Treaty Searchable Database* provides the following ERK and ERM functionalities, but without the application of conventional metadata (see questions 4, 4d and 5a) or folders for organizing the digital record entities:

Record Declaration

- Assign unique identifiers to digital record entities (Box 2).

Record Capture

- Allow import of digital record entities from other sources.

Record Organization

- Allow implementation of an agency-specific scheme to organize digital records entities;
- Allow selection of hierarchal categories to define the relative locations of digital records in the overall collection;
- Allow digital record entities to be linked to other records;
- Allow creation, addition, editing or deletion of digital record categories to implement the hierarchies.

Record Security

- Prevent over-writing of a record by users;
- Prevent any modification of a record's unique identifier once it is defined;
- Prevent deletion of indexes, categories, and other 'pointers' to records by users;
- Calculate and maintain a checksum for records;
- Provide logs of retrieval activity;
- Maintain appropriate backup copies of records;
- Provide adequate recovery/rollback procedures and rebuild procedures, so that records may be recovered or restored following a system malfunction.

Record Access

- Control access so that only authorized individuals are able to retrieve, view, print, copy, or edit a record or other entities in the system;
- Permit the identification of individual users and groups of users, and enable different access privileges to be assigned to individuals or groups. Access privileges could limit access to selected records or groups of records as well as by selected individuals.

Record Retrieval

- Allow searching on record content (without a controlled vocabulary);
- Ensure that all access privileges (permissions and restrictions) are enforced on all retrievals (open-access currently);
- Allow searching based on a combination of header tags, content, and subject categories within a single query;
- Allow retrieval of digital record entities and header tags;
- Provide a sufficiently powerful range of search features and options, as needed to meet various agency requirements.

Record Preservation

- Ensure that all records can be read and accurately interpreted throughout their useful life in that system;
- Enable migration of records to new storage media or formats to avoid loss due to media decay or technology obsolescence;
- Monitor storage capacity and utilization and alert system operators when action is needed.

Additional ERK or ERM functionalities could be implemented (as described by the National Archives and Records Administration) to accomplish agency-specific directives.

18a. Do the recordkeeping system(s) (or processes) routinely capture all digital entities within the scope of the activity it covers?

Yes. Operation of the *Antarctic Treaty Searchable Database* provides comprehensive access to the digital record entities. This access is logged by the record-keeping activities, as described in question 18.

18b. From what applications do the recordkeeping system(s) inherit or capture the digital entities and the related metadata (e.g., email, tracking systems, workflow systems, office systems, databases, etc.)?

It is not necessary with the DIGIN[®] technologies and is not captured to implement the *Antarctic Treaty Searchable Database* (see questions 4, 4d, 5a and 18). After the initial implementation of the *Antarctic Treaty Searchable Database* in 1999, the only captured files are the entire Antarctic Treaty Consultative Meeting (ATCM) Final Reports without metadata that have been published on the ATCM Web sites of the host nations. The new “*measures*” that have been adopted by the Antarctic Treaty Consultative Parties are then extracted and added to the *Antarctic Treaty Searchable Database* with header tags that define their unique location in the overall collection (see questions 4d and 4f).

18c. Are the digital entities organized in a way that reflects the creation processes? What is the schema, if any, for organising the digital entities?

Yes. The “*measures*”, which are extracted from the Antarctic Treaty Consultative Meeting (ATCM) Final Reports (see question 18b), contain header tag and title information (Box 2) that directly reflects their creation by the Antarctic Treaty Consultative Parties. Organization of the digital record entities (i.e., “*measures*”) effectively occurs with the dynamic generation of the hierarchical displays, which express relationships of the digital record entities within and between digital record series based on the ordering of the categorical contents in the header tags. The collection of digital record entities itself, absent the integration step with the Integration Module (Fig. 6), has no organization.

18d. Does the recordkeeping system provide ready access to all relevant digital entities and related metadata?

Yes. The recordkeeping system, which is the DIGIN[®] implementation (Fig. 6) of the *Antarctic Treaty Searchable Database* itself, provides comprehensive integrated access to the digital record entities (see question 18). The *Antarctic Treaty Searchable Database* does not require metadata (see questions 4, 4d and 5a).

18e. Does the recordkeeping system document all actions/ transactions that take place in the system re: the digital entities? If so, what are the metadata captured?

Yes. All queries of the Web site version of the *Antarctic Treaty Searchable Database* are automatically logged (see question 18). Metadata are not captured (see questions 4, 4d and 5a).

19. How does the creator maintain its digital entities through technological change?

Technology change has had no impact on the maintenance of the digital record entities. This technology independence exists because the digital record entities are maintained in

ASCII format, without markup or metadata that otherwise would create technology dependencies tied to “structured information” standards and legacy implementations.

As indicated by Table 3, there has been an increase in the number of digital record entities in the *Antarctic Treaty Searchable Database* over time. In some case, missing images also have been added to the digital record entities. These content changes among the digital record entities in the *Antarctic Treaty Searchable Database*, however, were uninfluenced by any technological changes between 2000 and 2004.

Table 3. Digital record entities in the *Antarctic Treaty Searchable Database* through time

TABLE 3: DIGITAL RECORD ENTITIES IN THE ANTARCTIC TREATY SEARCHABLE DATABASE THROUGH TIME						
Year Produced	Edition	Coverage ¹	Digital record Entities		Images	
			Number	Size (MB)	Number	Size (MB)
2000	1st	1959-1999	608	2.65	113	2.19
2001	2nd	1959-1999	608	2.65	164	4.46
2002	3rd	1959-2002	661	3.11	166	5.07
2003	4th	1959-2003	720	3.67	200	6.67
2004	5th	1959-2004	740	5.60	224	9.57

NOTE: Archived digital files on the webCDserver (Editions 1-5) are replicates of those on the Web site (<http://aspire.tierit.com/>). Functionalities to integrate the digital files on the Web site and webCDserver have been unmodified since their initial implementation in 2000.

19a. What preservation strategies and/or methods are implemented and how?

The preservation strategies are described in question 18. The use of ASCII formats helps to ensure that the digital record entities can be accurately interpreted and migrated to new storage media or formats.

Throughout, the digital record entities in the *Antarctic Treaty Searchable Database* have been maintained on servers with back-up copies on additional hard-drives as well as on webCDservers (Table 3), which replicate the full functionality and contents of the Web site. In one instance, the webCDserver was used to restore the Web site for the *Antarctic Treaty Searchable Database*.

19b. Are these strategies or methods determined by the type of digital entities (in a technical sense) or by other criteria? If the latter, what criteria?

No. Original content of the digital record entities in the *Antarctic Treaty Searchable Database* is preserved along with their provenance information (Box 2). Maintaining the ASCII format ensures that all records can be read and accurately interpreted throughout their useful life in that system. The ASCII format also enables flexible migration of records to new storage media or formats in order to avoid loss due to media decay or technology obsolescence.

20. To what extent do policies, procedures, and standards currently control records creation, maintenance, preservation and use in the context of the creator's activity? Do these policies, procedures, and standards need to be modified or augmented?

The policies and procedures illustrated in the Activity-Flow-Diagram (Fig. 4) are essential to the implementation of the *Antarctic Treaty Searchable Database*, as defined by Rules 1-9. Although the DIGIN[®] technologies are interoperable with metadata, markup and databases—all of which have well-defined patterns that reveal rules for the operation of the Granularity Module (Fig. 6)—implementation of the *Antarctic Treaty Searchable Database* is not controlled by any standards (see questions 4, 4d, 5a and 18).

21. What legal, moral (e.g. control over artistic expression) or ethical obligations, concerns or issues exist regarding the creation, maintenance, preservation and use of the records in the context of the creator's activity?

There is a longstanding interest and activity on the part of the creator (Berkman 2002) to help ensure “*that Antarctica shall continue forever to be used exclusively for peaceful purposes...for the progress of all mankind,*” as stated in the preamble of the *1959 Antarctic Treaty*. However, there are no legal, moral or ethical obligations regarding the implementation of the *Antarctic Treaty Searchable Database* in either its Internet or webCDserver forms (see question 15). The responsibilities and interests of the creator are stated in the Web site disclaimer and *End Users License Agreement*, as described in question 10.

The *Antarctic Treaty Searchable Database* began evolving in the United States in 1999 (see question 15) with a six-month funding supplement from the National Science Foundation, Division of Undergraduate Education and Office of Polar Programs, at a time when there were no other digital versions of the policy documents from the Antarctic Treaty System. In the United States, this support was particularly important since Antarctic policy and programs are directed through the Office of Polar Programs at the National Science Foundation, which has had the national responsibility to maintain an “*active and influential presence in Antarctica*” since Presidential Memorandum 6646 was issued on February 5, 1982.

Since 1999, the United States Department of State also has been providing access to public-domain policy documents from the Antarctic Treaty System, which has been essential to continuously update the *Antarctic Treaty Searchable Database*. These ongoing contributions from the Department of State are particularly important because the United States is the depository government for the *1959 Antarctic Treaty*. With the initial support from the National Science Foundation and the Department of State, the *Antarctic Treaty Searchable Database* was produced in a manner that became helpful to the international community as indicated by its being linked to the Web sites for the 24th and 25th Antarctic Treaty Consultative Meetings in St. Petersburg and Warsaw, respectively in 2001 and 2002 (Box 5).

Since 2003, the *Antarctic Treaty Searchable Database* has been funded through the National Science Digital Library program at the National Science Foundation as part of a

larger project to implement the “*Marine Mammal Commission Digital Library of International Environmental and Ecosystem Policy Documents.*” Today, the *Antarctic Treaty Searchable Database* is being used at the international government level by the Antarctic Treaty Secretariat and at the national level by government agencies in several countries as well as by non-governmental organizations, corporations and academic institutions around the world (Table 2, Box 5).

The level of international acceptance of the *Antarctic Treaty Searchable Database* (Table 2, Box 5) was never anticipated and, although there is no obligation to continue its implementation, there is a deep sense of responsibility on the part of the creator to sustain its utility “*for the progress of all mankind.*” The challenge is learning how to best contribute technological innovations that could “*facilitate the exchange of information*” and “*promote international cooperation*” in the operation of the Antarctic Treaty System, particularly through the new Antarctic Treaty Secretariat. As we build our world information society (<http://www.itu.int/wsis/>), evolution of the *Antarctic Treaty Searchable Database Case Study* also may reveal lessons that apply to the operation of other legal regimes.

22. What descriptive or other metadata schema or standards are currently being used in the creation, maintenance, use and preservation of the recordkeeping system or environment being studied?

Descriptive metadata, as conventionally applied with templates and attributes that reside in repositories, are not used to implement the *Antarctic Treaty Searchable Database* (see questions 4, 4d, 5a, 18 and 20). The descriptive schemas are the complete content of the digital record entities (see questions 5a, 5b and 18) and their header tags that describe the parent-child provenance (Box 2).

However, conventional metadata regarding the portal for the *Antarctic Treaty Searchable Database* (<http://aspire.tierit.com/>) are being added to the National Science Digital Library (<http://www.nsdlib.org>) and Digital Library for Earth System Education (<http://www.dlese.org>). The metadata format for these submissions is a modified Dublin-Core metadata (<http://dublincore.org/>) with additional fields for the education audiences that are being addressed by these digital libraries.

23. What is the source of these descriptive or other metadata schema or standards (institutional convention, professional body, institutional convention, international standard, individual practice, etc.?)

Conventional metadata (as described in question 22) are unnecessary with the DIGIN[®] technologies, which can interoperate with or without metadata to integrate “structured” as well as “unstructured” information (see questions 4, 4d, 5a, 5b, 18, 20 and 22). Moreover, metadata are not used with the *Antarctic Treaty Searchable Database* to organize, preserve, comprehensively integrate (Figs. 2 and 3) or quantify features among (Figs. 7 and 10) the digital record entities. The sources of the descriptive schemas are the persistent digital record entities themselves.

CONCLUSION

InterPARES 2 has implemented case studies to interpret the “*dynamic, interactive and experiential components*” of digital records as opposed to hardcopy or analog records. In an archival context, traditionally, all records are fixed at the time of their preservation. This concept of fixity is in direct opposition to the notion of a dynamic record, which is inherently variable. This apparent contradiction between a fixed record and a dynamic record raises several questions:

1. Are there differences between dynamic records and dynamic processes in digital environments?
2. What are the aspects of fixity that are consistent between hardcopy and digital records?
3. How can fixity be re-defined in terms of digital records that are dynamically generated?

Answers to these questions rely on distinctions between digital and the hardcopy media that have been used to share written information within and between populations since the beginning of our civilization (Fig. 11).

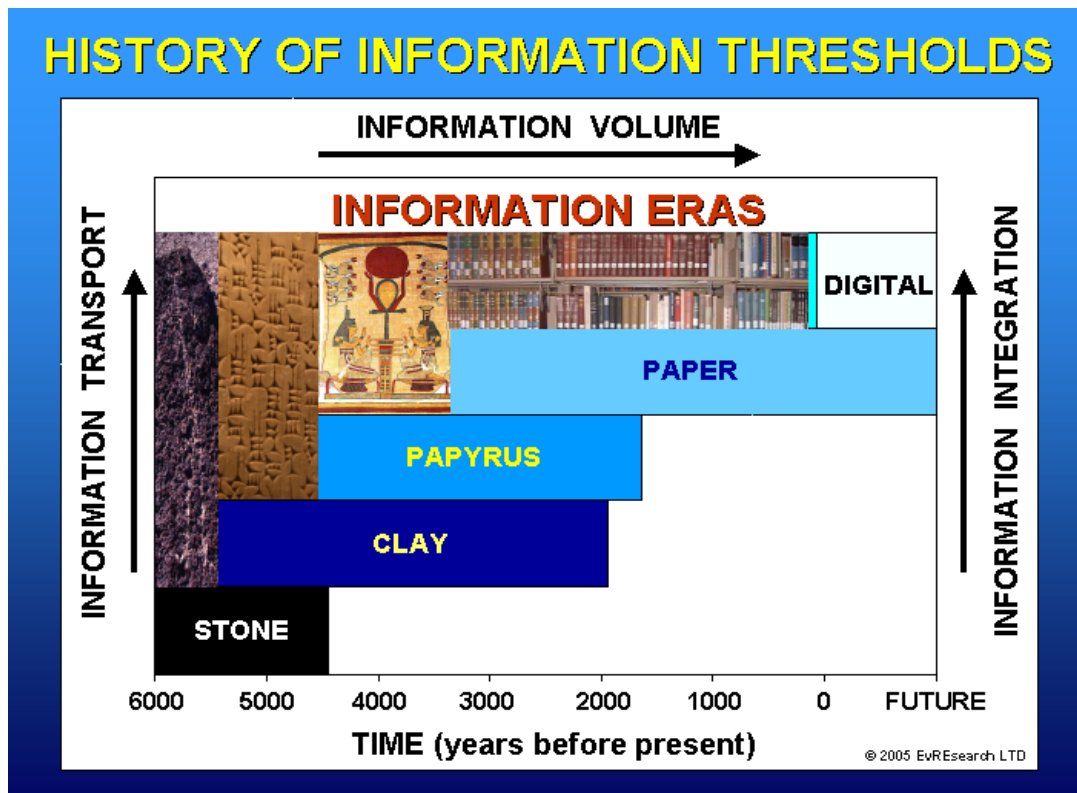


Figure 11. Thresholds in the preservation and dissemination of written information in our civilization. Starting arbitrarily at 6000 years ago, each of the media prior to digital had been used for millennia (Senner 1989). From stone to digital media: (1) the transport of information across time and space has increased; (2) the volume and rate of information produced has increased; and (3) the capacity to integrate information into new knowledge has increased.

The practical result of fixity in an archive is consistent and reproducible access to records, which is an absolute requirement. With the *Antarctic Treaty Searchable Database*, the dynamic integration of digital record entities results in consistent and reproducible hierarchies for any prescribed set of years (Figs. 2, 3, 7 and 10). When the *Antarctic Treaty Searchable Database* is updated with new information, it still is possible generate consistent and reproducible hierarchies across a prescribed set of years for any query,. Consequently, the results of an objective integration are fixed even though they are dynamically generated.

The basis for the objective integration in the *Antarctic Treaty Searchable Database* derives from the ability to automatically utilize the inherent structure of digital record groups to manage the underlying digital record series and digital record entities (Box 3). Structure, along with content and context, are the indivisible ingredients of information. Remove one of these ingredients and information has no meaning. For example, when a message is encrypted (i.e., the structure is altered) it still has content and context—but no meaning absent the key to decode the structure. A book can be managed based on its content (as in libraries) or its context (as in archives), but it is not possible to dynamically manage the sub-sections of any hardcopy media. Similarly, meaning is obscured if the dates, names or places (i.e., the content) are removed. If the authenticity of information (i.e., the context) is unknown, the validity of the information is questionable.

The paradigm shift created by digital technologies (e.g., Fig. 6) is the opportunity to **dynamically and objectively manage the structure of information** as well as the content and context. Unlike the subjective decisions that may vary from person to person to describe the context and content of a record, the structure is an inherent element of a record that can be described objectively. It is this ability to automatically utilize the inherent structure of information that distinguishes information management with digital media from the hardcopy media that had been applied previously in our civilization (Fig. 11).

Because integration results are objectively based on the inherent structure of the digital records (Box 3), as demonstrated by the *Antarctic Treaty Searchable Database*, dynamically -generated hierarchies (Figs. 2 and 3) can be used to produce quantitative results (Figs. 7 and 10) that are reliable and reproducible. Consequently, the concept of fixity is maintained by the relational schemas that result from dynamic integrations with the same set of granules for any query. In summary, the *Antarctic Treaty Searchable Database* reveals that dynamic records do exist and that they can be consistently reproduced in a digital environment.

E. References Cited

- Berkman, P.A. 1997. "The Zen of Venn," *Transactions of the American Geophysical Union (EOS)* 78(20):209.
- Berkman, P.A. 2002. *Science into Policy: Global Lessons from Antarctica*. Academic Press, San Diego.
- Bobak, A. 1997. *Data Modeling and Design for Today's Architectures*. Artech House Publishers, Norwood.
- Morgan, G.J. 2001. *Information Management, Retrieval and Display Systems and Associated Methods*. United States Patent and Trademark Office, Patent No. 6,175,830.
- Morgan, G.J. 2002. *Information Management, Retrieval and Display Systems and Associated Methods*. United States Patent and Trademark Office, Continuation Patent No. 6,484,166.
- Morgan, G.J. 2004a. *Information Management, Retrieval and Display Systems and Associated Methods*. Intellectual Property Office of New Zealand, Letters Patent No. 515007.
- Morgan, G.J. 2004b *Information Management, Retrieval and Display Systems and Associated Methods*. Commonwealth of Australia Patent Office, Letters Patent No. 770087.
- Senner, W.M. (ed.). 1989. *The Origins of Writing*. University of Nebraska Press, Lincoln.
- Song, J. and D. Whittington. 2004. "Why have some countries on international rivers been successful negotiating treaties? A global perspective," *Water Resources Research* 40 (W05S06):1-18.
- Yoffe, S., G. Fiske, M. Giordano, M. Giordana, K. Larson, K. Stahl and A.T. Wolf. 2004. "Geography of international water conflict and cooperation: data sets and applications," *Water Resources Research* 40(W05S04):1-8.

F. Glossary of Terms

AGGREGATION

Applies to related digital record groups, digital record series or digital record entities that are combined into a single file or folder based on objective or subjective determinations that meet the criteria of a user.

ARRANGEMENT¹

The body of principles and practices which archivists follow to group records in such a way as to reflect the manner in which they were held and used by the office or person creating the records. It involves the fundamental principles of *respect des fonds*, provenance, and sanctity of original order. The key units in archival arrangement are: record groups, sub-groups, and record series.

ARCHIVES²

Organized non-current records of an institution or organization retained for their continuing value in providing a) evidence of the existence, functions, and operations of the institution or organization that generated them, or b) other information on activities or persons affected by the organization. Derived from the Greek word for "government house," the term "archives" also refers to the agency responsible for selecting, preserving, and making available noncurrent records with long-term value and to the building or part of the building housing them.

ASCII

Acronym for the *American Standard Code for Information Interchange*. ASCII is a universal standard, based on string of seven binary digits, that defines how computers read and write Latin letters, numbers, punctuation marks and control instructions (<http://www.asciitable.com/>).

AUTHENTICITY³

Degree of confidence a user can have that the object is the same as that expected based on a prior reference or that it is what it purports to be

BYTE-OFFSET

The location of a digital record entity within a digital record series or digital record group measured from a starting point in bytes.

CONCEPT SPACE⁴

An interconnected, weighted network of terms (vocabularies) that represent a logical self-contained unit of information with a shared theme within the scope a broader knowledge space

¹ University of Illinois at Urbana-Champaign, University Archives (<http://web.library.uiuc.edu/ahx/define.htm>).

² Ibid.

³ National Library of Australia (<http://www.nla.gov.au/padi/topics/4.html>).

⁴ National Center for Supercomputing Applications (<http://archive.ncsa.uiuc.edu/SDG/IT94/Proceedings/Searching/chen/chenschatz2.html>).

CONTENT

An indivisible element of information, along with context and structure, that provides meaning. This element of information involves facts and concepts that can be interpreted with subjective descriptions that may vary between individuals or algorithms.

CONTEXT

An indivisible element of information, along with content and structure, that provides meaning. This element of information involves relationships within the record as well as the authenticity of the record in relation to its creator and creation environment.

CREATOR⁵

A person, corporate body, or family who created, accumulated and used archival records in the conduct of personal or business life.

DATABASE⁶

A collection of related information about a subject organized in a useful manner that provides a base or foundation for procedures such as retrieving information, drawing conclusions, and making decisions.

DIGITAL RECORD

Applies to the digital versions of record groups, record series and record entities, as defined below.

DOCUMENTS⁷

are instruments for the communication of information, regardless of their physical form or characteristics. They may be in the form of an impression on paper, a magnetic impulse, or a beam of light. The word comes from the Latin for official paper or that which teaches. Essentially, documents provide evidence or support of an action, condition, or entity.

DYNAMIC RECORD

The consistent and reproducible product of a digital integration of granules (see Granules, below).

ELECTRONIC RECORD KEEPING SYSTEM⁸

An electronic system in which records are collected, organized, and categorized to facilitate their preservation, retrieval, use, and disposition (Source: 36 CFR 1234.2). An electronic recordkeeping system may be either a distinct system designed specifically to provide recordkeeping functionality or part of another system. A distinct electronic recordkeeping system will comprise an application program which provides

⁵ University of Toronto, Archival Information Network (ARCHAEION) (<http://archeion-ao.fis.utoronto.ca/glossary.html>).

⁶ American Society of Crime Laboratory Directors / Laboratory Accreditation Board (<http://www.ascd-lab.org/aslab022.html>).

⁷ University of Illinois at Urbana-Champaign, University Archives (<http://web.library.uiuc.edu/ahx/define.htm>).

⁸ National Archives and Records Administration

http://www.archives.gov/records_management/policy_and_guidance/print_friendly.html?page=bulletin_99_04_a_content.html&title=NARA%20%7C%20Records%20Management%20%7C%20NARA%20Bulletin%2099-04%20A; Code of Federal Regulations (<http://www.gpoaccess.gov/cfr/index.html>).

recordkeeping functionality, data and metadata needed for management of the records controlled by the system, and any electronic records managed by the system. An electronic recordkeeping system may be part of another system, such as an application system or an electronic document management system, when the design of that system includes recordkeeping functionality.

FIXITY

The preservation quality of being immutable and unchanged over time or space. Result of the archival process that enables reliable and reproducible access to information with consistent content, context and structural elements.

FONDS⁹

The whole of the records, regardless of form or medium, automatically and organically accumulated by a particular individual, family, or corporate body in the course of that creator's activities or functions.

GRANULES

Granules are concept spaces within knowledge spaces that may be any size. Granules are analogous to record entities, which are logical units of information within record series or record groups. For the purposes of integration, granules reflect the smallest units in a "hierarchy."

GRANULARITY

The extent to which a system contains separate components (granules). Greater granularity refers to more identified components in the system. In an information system, the capacity to integrate information is proportional to the granularity.

HIERARCHY

Multidimensional display of relationships, which is contrasted with the linear displays of lists that hide relationships within and between digital records.

INTEGRATION

Creation of links between previously separate computer systems, applications, services, processes or components. Computer integration commonly applies to infrastructure of software and hardware components. Information integration applies to the creation or identification of relationships among digital record entities.

KNOWLEDGE DISCOVERY¹⁰

Interdisciplinary field which merges database management, statistics, machine learning and related areas to identify, extract and discover meaningful relationships or patterns from large collections of data. Discovery is directly related to granularity and integration.

KNOWLEDGE SPACE

Collection of feasible concept spaces or knowledge states that have structural connections within a domain.

⁹ University of Toronto, Archival Information Network (ARCHAEION) (<http://archeion-ao.fis.utoronto.ca/glossary.html>).

¹⁰ United Nations Educational and Cultural Organization (<http://www.netnam.vn/unescocourse/knowledge/1.1.htm>).

MARKUP

Detailed stylistic instructions within a digital record entity, digital record series or digital record group to organize units of information so that they can be read, printed and searched individually. Commonly-used markup languages are extensible markup language (XML), standard generalized markup language (SGML) and hyper-text markup language (HTML).

METADATA

“Data about data” that describes how, when and by whom the particular set of data was collected and how the data is formatted. Metadata commonly follows standardized formats, such as Dublin-Core Metadata (<http://dublincore.org>), to describe the attributes of the data. In addition, metadata commonly are harvested from the data and then stored in repositories for subsequent access by search engines.

PERL

Practical Extraction and Reporting Language is a stable cross-platform programming language that can read and write binary files as well as process very large files.

PERSISTENCE

Propagation of authentic digital records and their functionalities over time, independent of hardware or software platforms.

PROVENANCE¹¹

The principle of archival arrangement according to which each deposit of records should be placed within an overall arrangement or classification scheme that reflects its origin and relation to other deposits from the same administrative body.

RECORDS¹²

All documents, regardless of form, produced or received by any agency, officer, or employee of an institution or organization in the conduct of its business. Documents include all forms of recorded information, such as: correspondence, computer data, files, financial statements, manuscripts, moving images, publications, photographs, sound recordings, drawings, or other material bearing upon the activities and functions of the institution or organization, its officers, and employees. A document becomes a record when it is placed in an organized filing system for use as evidence or information. It becomes archival when transferred to a repository for preservation and research use.

*RESPECT DES FONDS*¹³

The principle of archival arrangement according to which each deposit (*fonds*) should be maintained as a separate entity, even if other *fonds* cover the same or similar subjects. It requires archivists to respect the integrity of the body of records at the time it is deposited in the archives.

¹¹ University of Illinois at Urbana-Champaign, University Archives (<http://web.library.uiuc.edu/ahx/define.htm>).

¹² Ibid.

¹³ Ibid.

RECORD ENTITY

Distinct logical unit of information (granule) that reflects the granularity in a knowledge space. Record entities can be derived from larger units of information as well as aggregated into a composite of information components.

RECORD GROUP¹⁴

A body of organizationally related records, normally large in size and established on the basis of provenance to accommodate the records of major organizational units and functions of an institution.

RECORD SUB-GROUPS¹⁵

Smaller (than record groups) bodies of organizationally related records placed within a record group to correspond to the subordinate administrative units that collectively form the record group.

RECORD SERIES¹⁶

Series means file units or documents arranged according to a filing system or kept together because they relate to a particular subject or function, result from the same activity, document a specific kind of transaction, take a particular physical form, or have some other relationship arising out of their creation, receipt, or use, such as restrictions on access and use (36 CFR 1220.14).

RULE SET

A set of rules, definitions or instructions for the operation of computer systems and programs.

SANCTITY OF THE ORIGINAL ORDER¹⁷

The principle of archival arrangement according to which the creator's arrangement of files and documents within a deposit should be maintained.

STRUCTURE

An indivisible element of information, along with content and context, that provides meaning. This element of information involves the inherent organization of the record with patterns, formats and arrangements of content segments or granules that can be objectively expressed in terms of parent-child relationships.

STRUCTURED INFORMATION

Information that is manipulated with databases, metadata or markup to express organizational or conceptual features that have been identified by a user based on the content of a digital record. An artificial class of digital records, which is contrasted with unstructured information, that exists because relational schema are derived without applying the inherent structures within digital records.

¹⁴ Ibid.

¹⁵ Ibid.

¹⁶ Code of Federal Regulations (<http://www.gpoaccess.gov/cfr/index.html>).

¹⁷ University of Illinois at Urbana-Champaign, University Archives (<http://web.library.uiuc.edu/ahx/define.htm>).

UNSTRUCTURED INFORMATION

Conventionally considered to be information that cannot be decomposed into relational schema, such as: pictures in any format, written documents or multimedia content as audio and video files. The term is a misnomer since all information has both content and inherent structure. For example, an encrypted message has content, but with a hidden structure that obscures any meaning. Moreover, the *Antarctic Treaty Searchable Database* (<http://aspire.tierit.com/>) and other DIGIN[®] databases demonstrate that digital records can be automatically decomposed into relational schema without metadata, markup or other imposed structural formats.

USE¹⁸

The retrieval of information from archival and manuscript holdings, finding aids, reference tools, and staff memories. Regardless of purpose, such as administrative action, publication of a book, preparation of a course paper, genealogy, or personal curiosity, any retrieval of information should be considered as use.

WebCDserver¹⁹

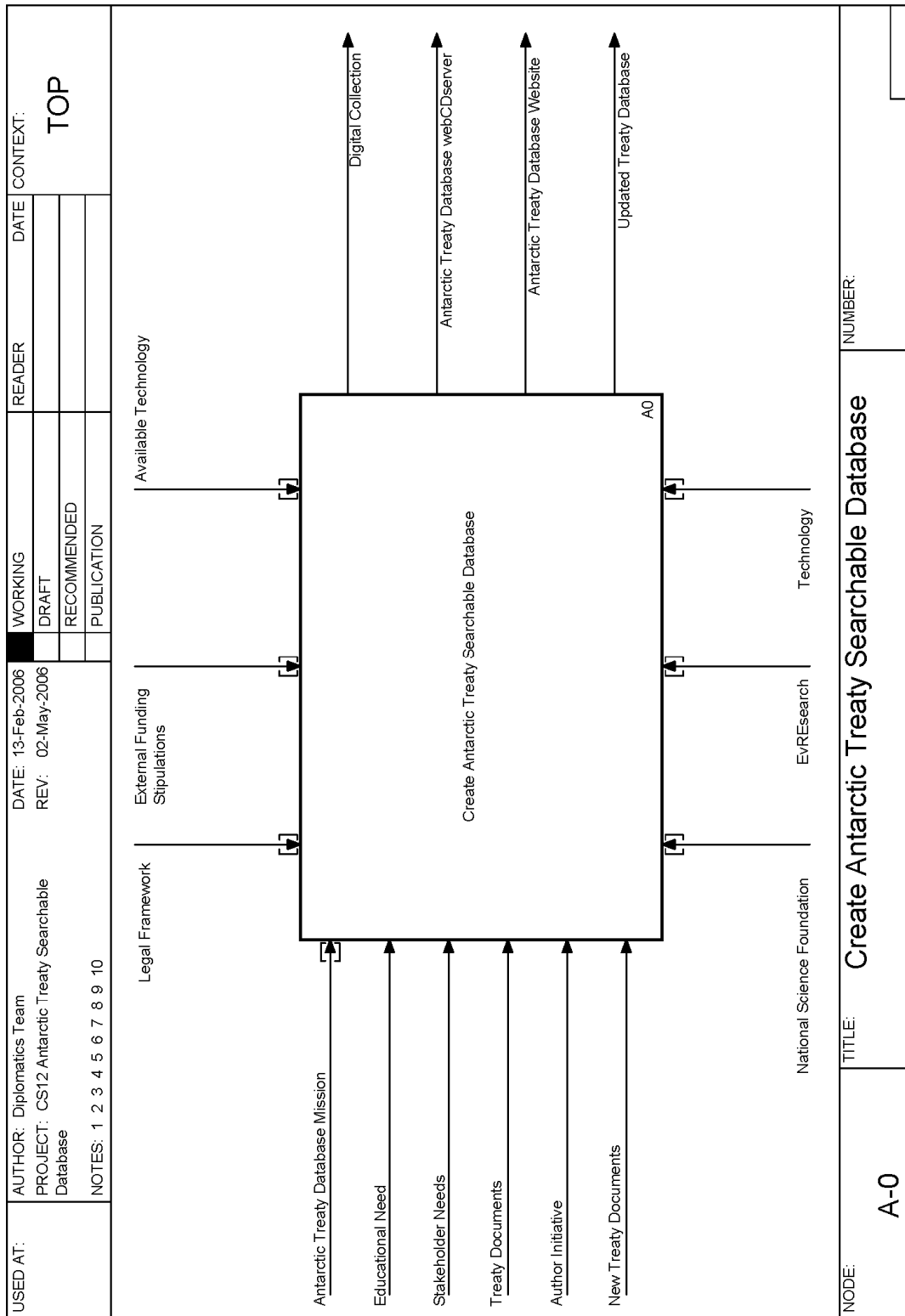
World-Wide-Web servers that can be run from CD-ROM, DVD or other static storage device without requiring installation to dynamic storage devices such as computer hard drives.

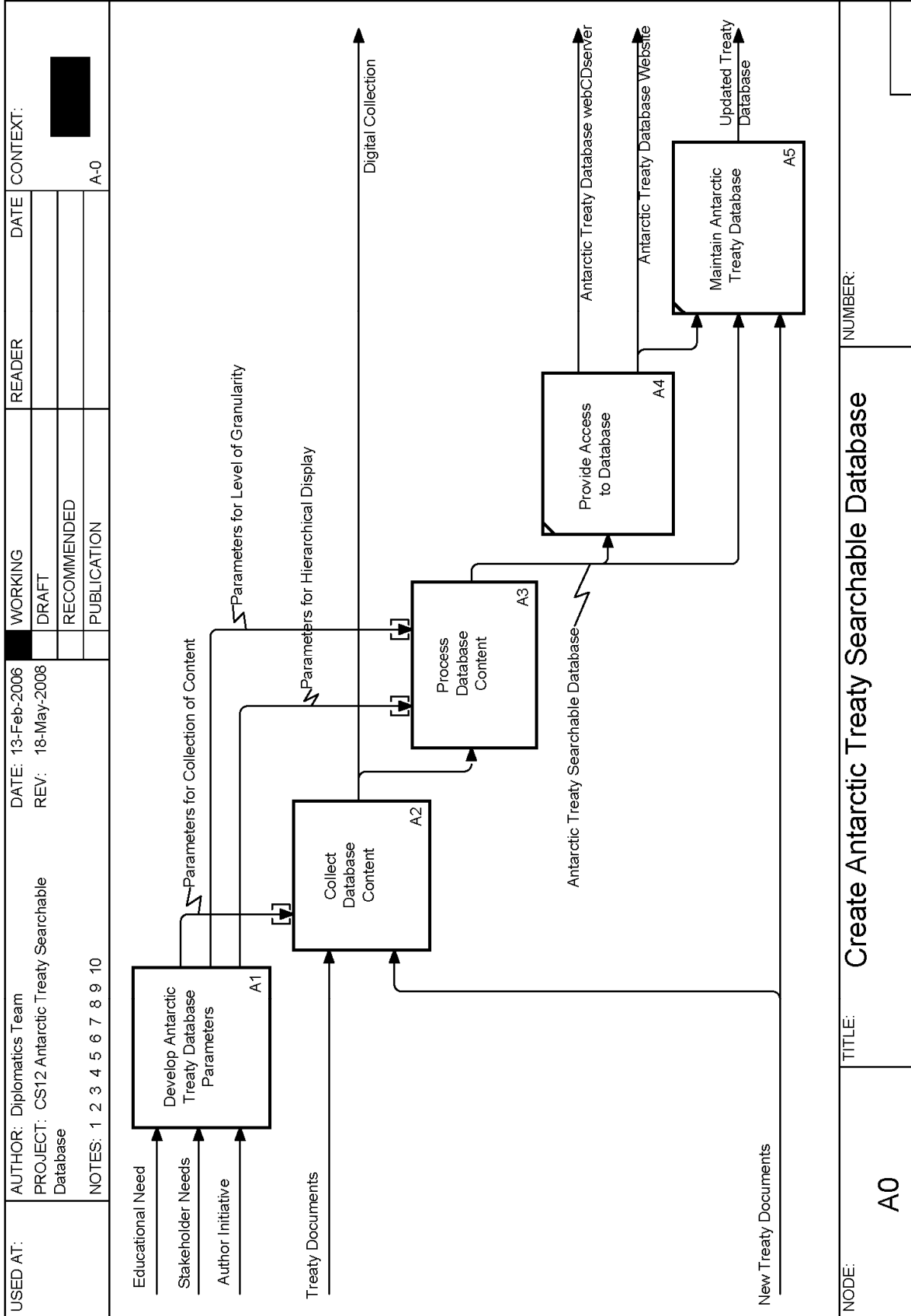
NOTE: There is an additional glossary from InterPARES 1, which has different definitions for some of the above terms. The above terms are being proposed for consideration to supplement the scope of the definitions in the searchable glossary and controlled vocabulary that are being constructed by the Terminology Cross-domain Task Force in InterPARES 2 (Fig. 1).

¹⁸ Ibid.

¹⁹ EvREsearch LTD (<http://www.evresearch.com>).

G. IDEF0 Activity Model



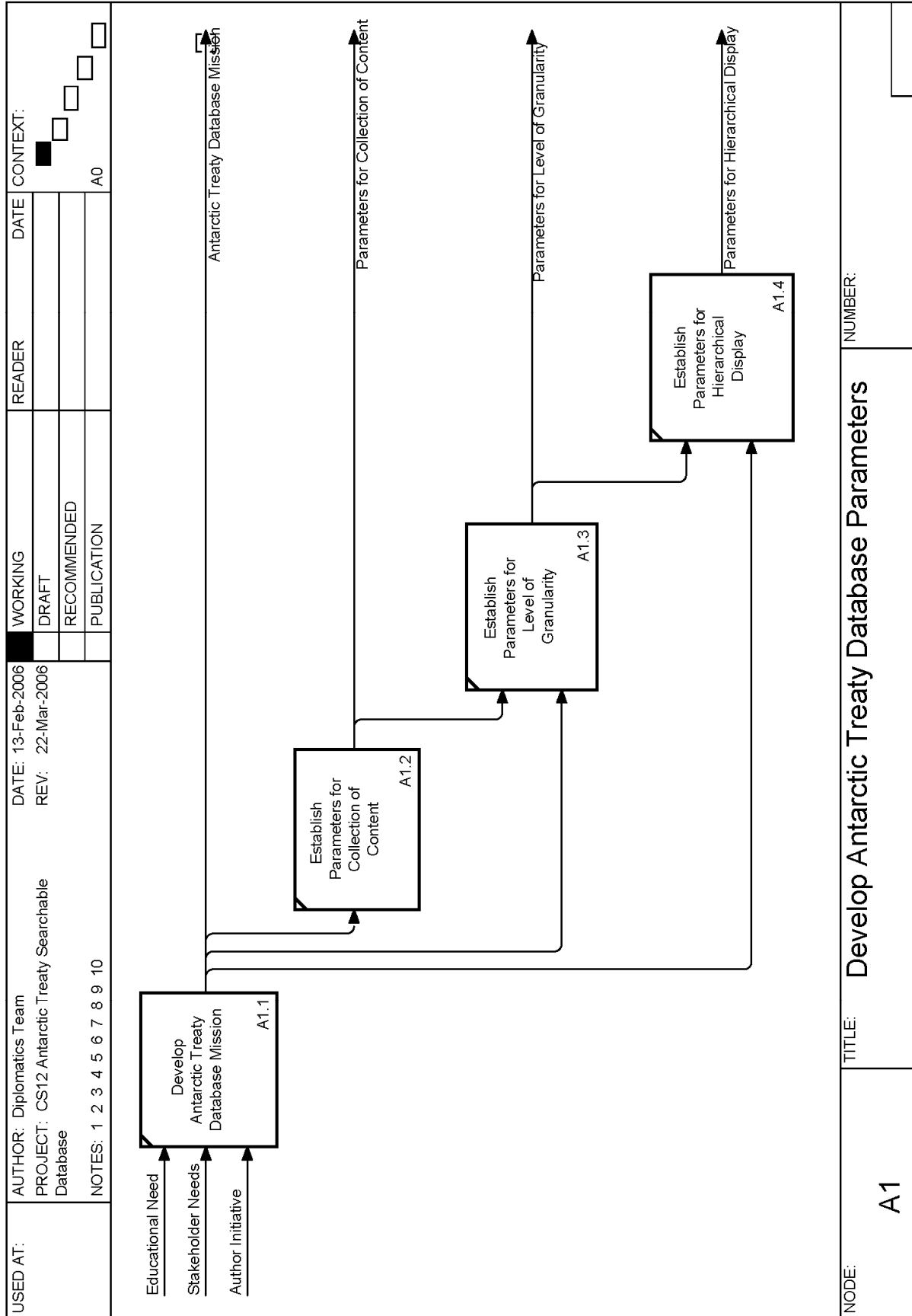


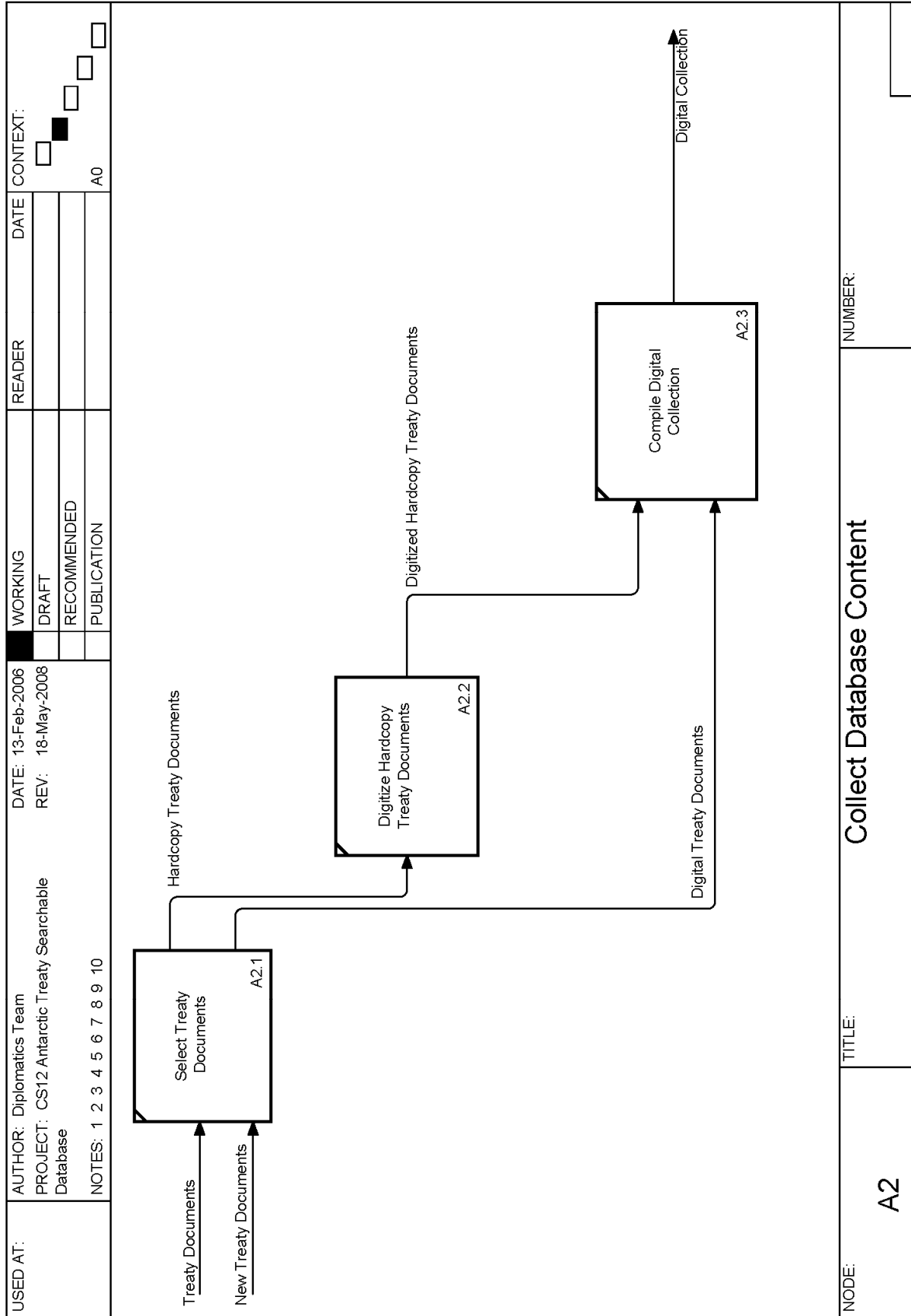
NUMBER:

Create Antarctic Treaty Searchable Database

TITLE:

A0





CS12 – Antarctic Treaty Searchable Database, IDEF0 Model: Activity Definitions (20060428)			
Activity Name	Activity No.	Activity Definition	Activity Note
Create Antarctic Treaty Searchable Database	A0	To develop Antarctic Treaty Database parameters, to collect database content, to process database content, to provide access to database, and to maintain Antarctic Treaty Database.	
Develop Antarctic Treaty Database Parameters	A1	To develop Antarctic Treaty Database mission, to establish parameters for collection of content, to establish parameters for levels of granularity, and to establish parameters for hierarchical display.	
Develop Antarctic Treaty Database Mission	A1.1	To interpret educational and stakeholder needs into database planning and design.	
Establish Parameters for Collection of Content	A1.2	To determine the scope of collection and compilation of database content, based on the Antarctic Treaty Handbook 8th Edition.	
Establish Parameters for Level of Granularity	A1.3	To identify the rules for the appropriate level of granularity of documents that will be searchable.	For example, each decision, measure or resolution is treated as a complete information granule. See p. 12 of the Final Report.
Establish Parameters for Hierarchical Display	A1.4	To determine the parent-child relationships between information granules.	See p. 12 of Final Report.
Collect Database Content	A2	To select Treaty documents, digitize hardcopy treaty documents and compile digital collection.	
Select Treaty Documents	A2.1	To choose treaty documents to be included in the database.	
Digitize Hardcopy Treaty Documents	A2.2	To convert selected treaty documents in hardcopy format to a digital format.	
Compile Digital Collection	A2.3	To combine all digital and digitized treaty documents to form database collection.	
Process Database Content	A3	To create granules, tag granules, index granules and assess granule relationships.	
Create Granules	A3.1	To identify information granules based on established parameters.	This activity is automatically completed by the DIGIN software used to create the database.
Tag Granules	A3.2	To add tags that allow the indexing and hierarchical display of the information granules.	This activity is automatically completed by the DIGIN software used to create the database, based on parameters established at the development stage.

Index Granules	A3.3	To structure the granule tags to facilitate hierarchical display.	This activity is automatically completed by the DIGIN software used to create the database, based on parameters established at the development stage.
Assess Granule Relationships	A3.4	To determine whether relationships between indexed tagged information granules conform to established parameters and meet user needs.	The database is automatically generated when tagged granules have been indexed. When the relationships have been assessed, the database can be considered to be complete and effective.
Provide Access to Database	A4	To provide integrated access to the Antarctic Treaty Searchable Database by making it available online and on CD-ROM.	
Maintain Antarctic Treaty Database	A5	To monitor the use and performance of the database, and to update and revise the database as needed.	

CS12 – Antarctic Treaty Searchable Database, IDEF0 Model: Arrow Definitions (20060428)		
Arrow Name	Arrow Definition	Arrow Note
Antarctic Treaty Database Mission	The goals and objectives to be reached by the creation of the Antarctic Treaty Database, based on perceived educational and stakeholder needs.	
Antarctic Treaty Database webCDserver	The Antarctic Treaty Database available on webCDserver.	
Antarctic Treaty Database Website	The Antarctic Treaty Database available on the Internet.	
Antarctic Treaty Searchable Database	The database created to provide efficient and comprehensive access to Antarctic Treaty documents.	
Author Initiative	The author's response to the perceived lack of organized access to Antarctic treaties.	The author contacted the U.S. Department of State about their plans for digitizing treaty documents. Because they had no plans to do so, the author decided to create a tool to provide digital access to the documents. See p. 8 of Final Report.
Available Technology	The state of technology available at the time of the conception and creation of the Antarctic Treaty Database.	
Digital Collection	The collection of digital and digitized treaty documents that provide the content of the database.	
Digital Treaty Documents	Selected treaty documents in digital format.	
Digitized Hardcopy Treaty Documents	Selected hardcopy treaty documents that have been converted to digital format.	
Educational Need	The need perceived by the creator for better access to Antarctic Treaty documents for educational purposes.	The creator designed the database to facilitate knowledge discovery about the policies and strategies related to Antarctic treaties. See p. 11 of Final Report.
EvREsearch	The company that created the software used to build the database.	The company consists of the database's two creators.
External Funding Stipulations	Limitation and regulations imposed upon the creators by external funders.	The National Science Foundation has issued a number of grants to the creators.
Hardcopy Treaty Documents	Selected treaty documents in hardcopy format.	
Indexed Tagged Information Granules	Information granules that have been tagged and indexed.	

Information Granules	The smallest conceptual units that can be resolved directly from the overall collection.	
Legal Framework	Legislation such as US copyright and patent laws that affects the creation and maintenance of the Antarctic Treaty Database.	
National Science Foundation	An external agency providing funding for the Antarctic Treaty Database project.	
New Treaty Documents	New documents related to Antarctic treaties and policies.	New documents are added to the database as they become available.
Parameters for Collection of Content	The rules and limits for the collection and compilation of database content.	Include measures, tables and figures, and exclude extracts and introductory notes. See p. 13-14 of Final Report.
Parameters for Hierarchical Display	The rules for hierarchical display of information granules.	
Parameters for Level of Granularity	The rules defining appropriate levels of granularity.	
Stakeholder Needs	The needs of stakeholders in government, education, industry and the general public for the efficient retrieval of accurate and reliable information.	
Tagged Information Granules	Information granules with attached tags.	
Technology	The DIGIN software used to create the database, as well as the hardware and software used to disseminate it.	DIGIN software was developed specifically for this project, and was patented by EvREsearch.
Treaty Documents	Documents related to Antarctic treaties and policies.	Many of these are found in the Antarctic Treaty Handbook, 8th Edition, but they may also come from other sources.
Updated Treaty Database	The revised Antarctic Treaty Database.	The Database will be revised as information is added to it.

Appendix 1: Abstracts

Open Science Conference

Scientific Committee on Antarctic Research (SCAR) XXVIII
Bremen, Germany, 25-31 July 2004

ANTARCTIC TREATY SEARCHABLE DATABASE - 1959 TO THE PRESENT

Paul Arthur Berkman

Bren School of Environmental Science and Management
University of California, Santa Barbara, CA 93106

ABSTRACT

The *Antarctic Treaty Searchable Database* (<http://webhost.nvi.net/aspire>) was initiated in 1999 as the first comprehensive searchable database of the international measures that have been adopted by the Antarctic Treaty Consultative Parties *in furtherance of the principles and objectives of the Treaty*. Conceived for classroom applications (as described in: Berkman, P.A. 2002. *Science into Policy: Global Lessons from Antarctica*. Academic Press, San Diego), the *Antarctic Treaty Searchable Database* soon became linked to: government agencies (e.g., Australian Antarctic Division, Environment Canada and National Academy of Sciences); non-governmental organizations (e.g., Scientific Committee on Antarctic Research and Antarctic Southern Ocean Coalition) and business (e.g., International Association of Antarctic Tour Operators) as well as academic institutions (e.g., George Washington University, Universität Freiburg and University of Canterbury). Moreover, the *Antarctic Treaty Searchable Database* was linked to the Web sites for the XXIV (2001) and XXV (2002) Antarctic Treaty Consultative Meetings in St. Petersburg, Russia, and Warsaw, Poland, respectively. In addition, the *Antarctic Treaty Searchable Database* is being utilized as an international case study for the *InterPARES (International Research on Permanent Authentic Records in Electronic Systems)* program that involves the national archives from 14 countries.

In contrast to conventional databases that generate long lists of ‘hits,’ the *Antarctic Treaty Searchable Database* dynamically generates expandable-collapsible hierarchies that comprehensively describe objective relationships among the relevant information resources for any search query. These information resources include the Antarctic Treaty as well as all of the recommendations, conventions, measures, decisions, resolutions, annexes, appendices, tables, figures and Protocol that have been approved by the Antarctic Treaty Consultative Parties from 1959 to the present. The underlying technology for integrating these Antarctic Treaty System documents is based on the EvREsearch® *Digital Integration System*™ (DIGIN®). This recently-patented technology, which facilitates knowledge discovery with automated granularity, also is being utilized by the National Science Digital Library program to integrate policy documents from other international legal systems and by the National Archives and Records Administration in applications research to implement a persistent archive of authentic digital records for the federal government of the United States.

Activity and data flow diagrams have been constructed to illustrate the ongoing implementation of the *Antarctic Treaty Searchable Database*, which is in its 4th Edition. The *Antarctic Treaty Searchable Database* will be demonstrated to illustrate its applications for science, education, government and public programs on a global scale. EvREsearch LTD is implementing the *Antarctic Treaty Searchable Database* with support from the National Science Foundation in collaboration with the Marine Mammal Commission and the United States Department of State.

19th International CODATA Conference
Council on Data for Science and Technology
The Information Society: New Horizons for Science
Berlin, Germany, 7-10 November 2004

ANTARCTIC TREATY SEARCHABLE DATABASE - CASE STUDY

Paul Arthur Berkman, Ph.D.
Bren School of Environmental Science and Management
University of California, Santa Barbara, CA 93106

and

EvREsearch LTD
1611 Tennyson Court, Columbus, OH 43235

ABSTRACT

The first searchable digital library of Antarctic Treaty documents was produced in early 2000 with access to policy documents through the US Department of State and funding from the National Science Foundation. The *Antarctic Treaty Searchable Database*, which is available online (<http://webhost.nvi.net/aspire>) as well as on webCDserverSM, is in its 5th Edition with all of the “*measures that have been adopted in furtherance of the principals and objectives of the Treaty*” from 1961 through 2004. Originally intended as a supplement for a university course on Antarctic science and policy, the *Antarctic Treaty Searchable Database* has become linked to Web sites from: international government institutions; national government agencies; non-governmental organizations; corporations; and academic institutions. In addition to being the first searchable digital library of Antarctic Treaty documents ever produced, the *Antarctic Treaty Searchable Database* has nearly 750 granules and is the most comprehensive source for automatically integrating information from the Antarctic Treaty System. The *Antarctic Treaty Searchable Database* dynamically generates hierarchical displays that comprehensively describe objective relationships within and between policy documents based on the parent-child structure and contents of each information granule. This interoperable knowledge-discovery application is facilitated by automated granularity. Activity and data flow diagrams will be presented to describe the ongoing implementation of the *Antarctic Treaty Searchable Database* as a potential strategy for the dynamic exchange and integration of accurate, reliable and authentic digital records through the Antarctic Treaty Secretariat.

Session: “**Data Archiving: The InterPARES Project**” being held on Monday, November 8, 2004, from 1:45 P.M. to 3:15 P.M.

National Science Foundation / National Science Digital Library (NSDL)
Annual Meeting 2004
Chicago, Illinois, 14-17 November 2004

**AUTOMATED GRANULARITY AND COMPREHENSIVE INTEGRATION FOR
KNOWLEDGE DISCOVERY: A CASE STUDY WITH POLICY DOCUMENTS**

Paul Berkman (New Media Studio / University of California, Santa Barbara / EvREsearch LTD)
Howard Burrows (National Science Digital Library)
Bruce Caron (New Media Studio)
Julie Ekstrom (University of California, Santa Barbara)
Suzanne Montgomery (Marine Mammal Commission)
Reagan Moore (San Diego Supercomputer Center)
George Morgan (EvREsearch LTD, Native Voices International)
Oran Young (University of California, Santa Barbara)

ABSTRACT

This National Science Digital Library project (<http://nsdl.tierit.com>) is based on international environmental and ecosystem policy documents that were compiled in five volumes by the Marine Mammal Commission. One application of this NSDL collections project is the *Antarctic Treaty Searchable Database* (<http://webhost.nvi.net/aspire>), which is now in its 5th Edition with all of the policy documents that have been adopted by 45 nations in the Antarctic Treaty System from 1959 through 2004. Conceived for classroom use (as described in: Berkman, P.A. 2002. *Science into Policy: Global Lessons from Antarctica*. Academic Press, San Diego), the *Antarctic Treaty Searchable Database* is now linked to the Web sites of government agencies, non-governmental organizations, businesses and academic institutions around the world. Moreover, the *Antarctic Treaty Searchable Database* was linked to the Web sites for the 2001 and 2002 Antarctic Treaty Consultative Meetings in St. Petersburg, Russia, and Warsaw, Poland, respectively. In addition, the *Antarctic Treaty Searchable Database* is being utilized as an international case study for the *InterPARES (International Research on Permanent Authentic Records in Electronic Systems)* program that involves the national archives from 14 countries.

The underlying implementation of this NSDL collection is based on the patented EvREsearch[®] *Digital Integration System*[™] (DIGIN[®]) that:

- dynamically generates expandable-collapsible hierarchies that objectively identify relationships within and between information resources for any search query; and
- automatically turns qualitative information into quantitative data that can be used to interpret trends and facilitate decision-making.

DIGIN[®] operates by utilizing the inherent structure or patterns in digital information without requiring an understanding of the content to automatically create, comprehensively index and objectively integrate sets of information (“automated granularity”). The value of DIGIN[®] is that it uniquely provides both convenience and control to objectively and comprehensively manage digital information at scale. In contrast, search engines (e.g., Google[®]) are extremely convenient, but they provide limited control for users to manage or utilize all of the search results. The search-engine problem is that accessing **more information does not equal more knowledge**. Similarly, databases, metadata and markup provide control to manage digital information, but their applications are not convenient. Moreover, these conventional technologies do not work at scale, which is why less than 20% of the available digital information is “structured” and more than 80% is “unstructured.” The challenge, now and into the distant future, is to integrate 100% of the digital information based on user-defined objectives to discover knowledge and DIGIN[®] is a powerful innovation.