



InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

Overview

Case Study 12: Antarctic Treaty Searchable Database

Peter Gagné, Université Laval

May 2006

The Creator Context / Activity

Creator: EvREsearch Ltd.¹

Creator type: Scientific focus / Private sector (small corporation: two individuals, one of whom is involved in university research²). The creator is effectively functioning as an individual.

Juridical context: EvREsearch is a limited liability company that was founded in 1999 by two partners: Dr. Paul Arthur Berkman and George James Morgan III. Applicable legislation includes United States and international laws pertaining to copyright and patents. There are also laws and norms related to grants from the National Science Foundation.

It should be noted that while “the entire scope of the *Antarctic Treaty Searchable Database* is within the context of international law and other legal regimes,” (FR 27) it “has yet to be designed around any formal government mandate.” (FR 9)

Activity: The activity under study is the creation and maintenance of the Antarctic Treaty Searchable Database (ATSD),³ which allows the searching of official treaty documents, the integration of digital record entities and the dynamic generation of hierarchical displays to reveal relationships within and between digital records. Documents integrated into the database are from the collections of public domain policy documents that have been adopted by the Antarctic Treaty System from 1959 to 2004. This activity involves the use of EvREsearch’s patented Digital Integration (DIGIN[®]) System to automatically identify rules to automatically “break” the digital resources into discrete granules (e.g., articles within a convention).⁴

¹ See <http://www.evresearch.com/>.

² In addition to being a founder of EvREsearch, Dr. Berkman is also a research professor in the Bren School of Environmental Science & Management at the University of California, Santa Barbara.

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

⁴ DIGIN[®] technology and processes are described in detail in the final report on pages 15-18 and 24-25.

Nature of Partnership

The Antarctic Treaty Searchable Database was produced in early 2000 using materials from the United States Department of State. The initial information resource was the Antarctic Treaty Handbook, 8th Edition, published by the Department of State, which has continued to provide public domain policy documents. “These ongoing contributions from the Department of State are particularly important because the United States is the depository government for the 1959 Antarctic Treaty.” (FR 39) In addition to the handbook, contents for the ATSD came from other sources that were in either digital or hardcopy formats.

Continuing management and administration of the ATSD has been coordinated by Dr. Berkman, with input from an advisory board created in conjunction with his grant from the National Science Foundation, National Science Digital Library program. Additional input has been received from individuals in the National Science Foundation, Department of State and many other institutions in the U.S. and abroad that are involved with the Antarctic Treaty System. The ATSD is hosted by Native Voices International, which is separately owned by George Morgan, president of EvREsearch.

Initial funding was provided in 1998 by the National Science Foundation (Division of Undergraduate Education and Office of Polar Programs). Support to update the ATSD from 2000-2003 was provided by EvREsearch. Funding from the National Science Foundation (National Science Digital Library program) resumed from 2003-2006 in collaboration with the Marine Mammal Commission.

The ATSD also serves as an information source for a number of organizations, including international government institutions (Antarctic Treaty Secretariat); government agencies (Australian Antarctic Division, Environment Canada and the US Library of Congress); non-governmental organizations (Scientific Committee on Antarctic Research and Antarctic Southern Ocean Coalition); business (International Association of Antarctic Tour Operators) and academic institutions (George Washington University; Universität Freiburg, Oxford University and the University of California Santa Barbara).

Bureaucratic/Organizational Structure

EvREsearch Ltd has one location in Columbus, Ohio. Dr. Paul Arthur Berkman is the CEO and George James Morgan, III is the president of EvREsearch.

Digital Entities Studied

The digital entities studied are individual “information granules” or files that contain text documents (in ASCII format) with images (JPG files) and tables. The final report refers to the information granules as “digital record entities” and the ATSD itself as the “digital record.” The Granularity Module of the DIGIN[®] system was the principal tool for generating the digital entities that are the objects of this case study. Origin of the digital entities is from both digital and non-digital formats.

Documentary Practices Observed

Records Creation and Maintenance

The creator uses documented **procedures** for the creation of the digital entities. These processes are described in an Activity Flow diagram (FR 13) and a Data Flow diagram (FR 14).

“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.” (FR 22)

It is the **organization** of the information integrated into the ATSD that determines the organization of the digital entities. The latter mirror the structure of the documents from which the information is extracted, i.e., Year – Document – Measures – Article. As such, organization of the digital entities is “objective and defined by the inherent structure, patterns and organization of their parent digital record series.” (FR 24)

“The **aggregation** levels among digital entities in the *Antarctic Treaty Searchable Database* are based on the inherent parent-child relationships within the policy documents.” (FR 23)

Similar to the use of **metadata**, “granules” contain tags that define their unique parent-child relationships and which can be dynamically combined in relational schema. Implementation of EvREsearch’s patented DIGIN[®] System enables users to dynamically generate expandable/collapsible hierarchies that identify relationships within and between the relevant resources in the collection without metadata or markup.

Although DIGIN[®] is said to be interoperable with metadata, it seems that the creator has conceived this proprietary technology in **opposition to metadata**, which the creator claims is “is an inappropriate tool for user-defined levels of granularity beyond the digital record group and the digital record series.” (FR 25) There are two main differences between the use of tags in ATSD and the use of metadata: objectivity and integration. “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.” (FR 17) What is more, “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.” (FR 22)

With regards to **changes** made to the digital entities, Dr. Berkman informally updates and preserves data. “The digital record of the *Antarctic Treaty Searchable Database*, along with all of its elements (i.e., fonds), has been continuously updated...as: new ATCM “measures” are adopted by the parties 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.” (FR 14) However, “After they are implemented, content of the digital record entities is preserved and unmodified over time.” (FR 33)

Recordkeeping and Preservation

No formal records management or archives **program** exists. “The Antarctic Treaty Searchable Database operates without formal Electronic Recordkeeping or Electronic Records Management systems, as described by the National Archives and Records Administration.”⁵ (FR 35)

Maintenance and **preservation** responsibilities have been assumed “without formal mandate by Paul Berkman and EvREsearch LTD.” (FR 34) Moral, rather than archival, reasons motivate this preservation, seemingly bordering on the altruistic. “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.*’” (FR 40) Nonetheless, “it is hoped that the Antarctic Treaty Secretariat will take on these responsibilities to sustain the *Antarctic Treaty Searchable Database.*” (FR 34)

The current **preservation strategy** employed is to make backup copies of the database. Part of the activity under study includes the design and implementation of sustainability strategies for long-term preservation. This has led to the creation of webCDserver, a patent-pending technology to archive fully executable replicas of the Web site on static storage media. Each annual update of the database from 2000-2004 has been archived on a webCDserver.

The Web site automatically **logs** all queries (actions/transactions) to the ATSD.

Although the creator maintains that “technology change has had no impact on the maintenance of the digital record entities” (FR 37), various measures are used to avoid **obsolescence**. DIGIN[®] is written in PERL (Practical Extraction and Reporting Language), which provides a stable cross-platform programming language and the ASCII format is used, because it is a worldwide open-standard file format for representing Latin text, numbers, punctuations and symbols. This format allows for **migration** to new formats or media. In addition, “DIGIN[®] provides a general method that operates independently from any specific hardware and software...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.” (FR 15)

Accuracy, Authenticity and Reliability

The creator believes that questions of accuracy and reliability are answered by the utility of the ATSD and—to a great extent—the *perception* of accuracy and reliability by those who use it. “The *Antarctic Treaty Searchable Database* is accurate and reliable as reflected by its applications for diverse stakeholders around the world.” (FR 29)

Accuracy

The ATSD disavows any inaccurate data contained in the database. It clearly states, “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.” (FR 10)

⁵ See http://www.archives.gov/records_management/policy_and_guidance/examples_system_functions.html.

Authenticity

Authenticity of the ATSD, in the view of the creator, is equated with the authenticity of the source material integrated into the database. “The [elements of the database] have been ingested from authentic sources...and incorporated into the digital record of the *Antarctic Treaty Searchable Database* without modifying their content.” (FR 15) However, “authenticity, in the case of public-domain policy documents, can only be assured by the government agencies that issue the records,” (FR 29) and the only authentic documents relating to the Antarctic treaty, according to the creator, are the hardcopy versions produced by participating governments. According to this definition, the ATSD is not and cannot be authentic. Authenticity of the digital entities themselves throughout their lifecycle is not discussed in the final report.

Reliability

“To ensure reliability, the content and context of the Antarctic Treaty Searchable Database is preserved and unmodified from its original sources.” (FR 28) The creator believes that the reliability of the ATSD is shown by the fact that many international “stakeholders” have linked their Web sites to that of the ATSD.