InterPARES 2 Project
International Research on Permanent Authentic Records in Electronic Systems

Diplomatic Analysis

Case Study 06:
Cybercartographic Atlas of Antarctica

Sherry Xie, UBC

December 2006
(Revised February 2007)

INTRODUCTION

The InterPARES 2 case study 06, Cybercartographic Atlas of Antarctica (CAA), was conducted to examine the creation, maintenance, and preservation of an online interactive map system. It falls into InterPARES 2's investigation area of scientific activities that are carried out using experiential, interactive and dynamic computer technologies. The CAA project aims to bring together selected existing datasets in a dynamic fashion and to render them in a multimedia form, with emphasis on user interaction with geospatial information.

The following text presents the results of the diplomatic analysis on the digital entity identified in the case study report as study object. The purpose of the diplomatic analysis is to assess the status of the identified digital entity as record, and based on the analysis, Domain 3 of InterPARES 2 can propose applicable preservation strategies. The digital entity identified in this case study is the CAA. This diplomatic analysis therefore centres on the identification of the Atlas as a record.

The purpose of the diplomatic analysis is to assess the status of the identified digital entity as a record. Once the status of the digital entity has been determined, preservation strategies may be proposed by Domain 3.

IDENTIFICATION OF RECORD(S)

A record, as defined by the InterPARES glossary, is a document made or received and set aside in the course of a practical activity. A record must also possess all of the following five components, as established by InterPARES 1 research conclusions: fixed content and form, embedded action, archival bond, persons and contexts. The application of the definition of a record to the creator's (a) issue contents and (b) programming codes is therefore analyzed according to the following parameters:

1. To be identified as a record, the digital entity must possess fixed content and form,[1] and be affixed to a stable medium (or physical carrier).

- **The content of the Atlas is not fixed.**

  The following information extracted from the case study 06 final report indicates that the content of the CAA is not stable:

  - the answer to Q. 12 indicates, "[t]he CAA's creators[2] use some of the digital entities to continue to build and update the Atlas" (p. 15);
  - the answer to Q. 6 indicates "an Authors Toolkit is in production that will enable external, authorized content creators to add content directly to the online CAA" (p. 12);
  - after the content modules are integrated into the Atlas, the CAA's technical specialists can add or modify online content; and
  - the answer to Domain 1, Question 4, summarizes these processes in the following sentence, "the *CAA itself* is subject to ongoing, continuous content update, …" (p. 22, emphasis in original).

  While it is not clear whether the added new content modules overwrite the old ones that have the same theme,[3] the Atlas does not satisfy the requirement of stable content.[4] With such continuous updating, end users of the Atlas, who are promised to "search, follow hypertext links, construct a view of a map by selecting layers& features, time periods, themes, etc.," may experience different search results by the same inquiry or retrieve different texts follow the same hyperlink. The content of the Atlas is not stable but changing and growing.[5]

- **Whether the documentary form[6] of the Atlas is fixed cannot be decided at this stage.[7]**

---

[1] The InterPARES 1 Authenticity Task Force has defined fixed form as the following: 1) binary content of the record, including indicators of documentary form, must be stored in a manner that ensures it remains complete and unaltered, and 2) technology must be maintained and procedures defined and enforced to ensure that the content is presented or rendered with the same documentary form it had when set aside. (See ATF Research Methodology Statement, available at: http://www.interpares.org/documents/interpares_ResearchMethodologyStatement.pdf).

[2] Please note that the "creator" here is not the same "creator" used for the diplomatic analysis in this report.

[3] Content modules are developed based on specific themes.

[4] The concept of stable content refers to the fact that the data and the message in the record are unchanged and unchangeable. This implies that data cannot be overwritten, altered, deleted or added to. Thus, if one has a system that contains fluid, ever changing information, one has no records in such a system until one decides to make one and to save it with its unalterable content.

[5] This conclusion is purely based on the understanding gained from the descriptions in the final report since I do not have the opportunity to experiment with the Atlas.

[6] The InterPARES 1 Authenticity Task Force defines documentary form as: The rules of representation according to which the content of a record, its administrative and documentary context, and its authority are communicated. Documentary form possesses both extrinsic and intrinsic elements. Intrinsic elements are the discursive parts of the record that communicate both the action in which the record participates and the immediate context. The types of intrinsic elements include name of author, name of originator, chronological date, name of place of origin of record, name of addressee(s), name of receiver(s), indication of action (matter), and name of writer. Extrinsic elements refer to specific, perceivable features of the record that are instrumental in communicating and achieving the purpose for which it was created. The types of extrinsic elements include overall presentation (e.g., textual, graphic, image, sound, or combination of these), specific presentation features (e.g., special layouts, hyperlinks, colors), and special signs (e.g., watermarks, logo). The report of the task force is available at http://www.interpares.org/book/interpares_book_d_part1.pdf.

[7] Based on IP 2 findings: a digital entity has a fixed form when its binary content is stored so that the message it conveys can be rendered with the same presentation it had on the screen when first saved. Because the same presentation of a record can be produced by a variety of digital presentations, fixed form does not imply that the bit streams must remain intact over time.

The Atlas's documentary form is decided by both the intention of the content creator on how to display the content and the technologies (e.g., the compiler) used by the technology specialists to actualize the intention. Supposing the intention remains unchanged after the submission of the content module, the technologies used for its online presentation will be changing along with the availability of new technologies. The following information is extracted from the final report:

- After contents, which are marked using XML, are prepared by content module creators, technology specialists in the project process the contents via a compiler to make them Web-ready. "The compiler will change to meet the requirements of changing technology required to present the CAA on the Web" (Domain 1, Question 4, p. 22).
- "Should the technology platform of the CAA change, the content of the Atlas would be re-built by re-accessing the XML content modules and processing them anew through a new compiler" (Q. 17, p. 16).
- "[C]ontent modules can be re-compiled. The use of XML for the content modules should make the CAA easily translatable (via new compilers) into any future mark-up languages" (Q. 19, p 17).
- The answer to Domain 1, Question 4 summarizes these processes in the following sentence, "the *CAA itself* is subject to ongoing, continuous content update, as well as occasional technological upgrading" (p. 22, emphasis in original)
- The reason why fixed documentary form cannot be decided is because IP2 findings establish the concept of "bounded variability,"[8] which allows certain variations on the documentary form. It is unknown, however, at this stage what types of variations will be caused by new compliers.

There is no information found in the final report regarding where the CAA will be hosted since it is still under construction; however, it is reasonable to assume that it will be affixed to some server that acts as a physical carrier for the Atlas.

2. A record must also participate in an action, defined as the conscious exercise of will by an officer of the creator or by an external person, aimed to create, maintain, modify or extinguish situations. A record results as an unintended by-product or product of the action.

In the answer to Q. 2, *Which of these activities generate the digital entities that are the objects of your case study,* the final report identifies the action as "[t]he creation of the Cybercartographic Atlas of Antarctica (CAA)" (p. 6). According to this, the Atlas does not participate in the action and is not a by-product of such action; instead, it is the final product of the action.

[8] This concept refers to "changes to the form and/or content of a digital record that are limited and controlled by fixed rules, so that the same query, request, or interaction always generates the same result. Variations in the record's form and content are either caused by technology, such as different operating systems or applications used to access the document, or by the intention of the author or writer of the document. Where content is concerned, while, as mentioned, the same query will always return the same subset, its presentation might vary within an allowed range, such as image magnification. In consideration of the fact that what causes these variations also limits them, they are not considered to be violations of the requirements of fixed form and stable content (see Luciana Duranti and Kenneth Thibodeau (2006), "The Concept of Record in Interactive, Experiential and Dynamic Environments: the View of InterPARES," Archival Science 6(1): 13-68).

3. A record must possess an archival bond, which is the relationship that links each record to the previous and subsequent record of the same action and, incrementally, to all the records which participate in the same activity. The archival bond is originary (i.e., it comes into existence when a record is made or received and set aside), necessary (i.e., it exists for every record), and determined (i.e., it is characterized by the purpose of the record).

The Atlas does not possess an archival bond with other digital records generated in the process of carrying out the action, i.e., the construction of the Atlas.

There are many digital records generated in this action such as digital maps, layers, texts, audio/video files, games, content modules, etc., as well as technology constructions (e.g. codes captured in Subversion).[9] The relationship between the Atlas and these records, however, is not an archival bond but a composing relationship: the latter constitute the former. In other words, these digital records are the components of the Atlas. See below the detailed analysis on the relationship between the Atlas and other digital records in light of the three characteristics of archival bond:

a. *The relationship is not originary*: the Atlas will be hosted online once some of the content modules are created and technological preparations are done. Its existence on the Web and the continuity of construction indicate that it is not set aside[10] at this moment, and this makes it nowhere to acquire archival bond.

b. *The relationship is not necessary*: when the process of assembling various content modules and of deploying them on the Web has been completed (at one point over time, given the fact that new contents will be added to it), i.e., the Atlas comes into being, the its relationship with other digital records is not necessary since its meaning does not require the linkages to other digital records for being understood or interpreted.[11]

c. *The relationship is not determined*: the purpose of the Atlas is to fulfill the mandate of the Geomatics and Cartographic Research Centre (GCRC),[12] which is different from the purpose of other records, which are generated to produce an online Atlas.

[9] A diplomatic analysis will demonstrate they are records.
[10] The definition for "set aside" in the Glossary Definition in the InterPARES 2 Terminology Database is: "of a record that is filed or archived, by assigning to it a classification code, including it in a folder or consigning it to a recordkeeping system, or that is associated with other records by any other means."
[11] One may argue that these digital entities are still needed since, to answer user inquiries, the Atlas needs to search and organize digital entities that satisfy the query, thus without these individual entities, without the Atlas. Technically, this is true, given the dynamic and interactive nature of the online Atlas; but conceptually, once the Atlas ready for online launch, all of these digital entities become parts of the Atlas itself.
[12] The mandate of the GCRC is to increase knowledge related to human interaction with geospatial information. The Centre fulfils its mandate largely through research in cybercartography and high resolution remote sensing for surveying and mapping and natural resource applications. Genevieve Shepherd (with Tracey Lauriault), Characterization of Case Study, Case Study: CS06

4. Record creation must involve at least three persons, whether or not they explicitly appear in the record itself. These persons are the author, addressee and writer; in the electronic environment, one must also take into account two additional necessary persons: the creator and the originator.

- The record's **author** is the physical or juridical person having the authority and capacity to issue the record or in whose name or by whose command the record has been issued.

  The author of the Atlas is the Carleton University, a juridical person with the capacity to make the construction of the Atlas happen. While it has "numerous stakeholders," the CAA project is initialled as a key object of the Cybercartography and New Information Economy Project taking place in the Geomatics and Cartographic Research Center (GCRC) in the Department of Geography and Environmental Studies in the University.

- The **writer** is the physical or juridical person having the authority and capacity to articulate the content of the record.

  The writer of the Atlas is the CAA project team (including external content creators), who collectively produce the Atlas.

- The **addressee** the physical or juridical person(s) to whom the record is directed or for whom the record is intended.

  The addressee of the Atlas is the public. The Atlas will be electronically delivered via the Internet and the access is free.

- The **creator** is the person in whose fonds the record exists.

  If the Atlas is a record, it should belong to the fonds of the University, who is therefore the creator of the Atlas. In archival science, to organize records as a creator's fonds means these records were used for the creator's business purposes. However, the University does not create the Atlas for its own business purposes. As the answer to Q. 12, *How does the creator use the digital entities under examination?*, in the final report indicates:

  > The CAA's creators use some of the digital entities to continue to build and update the Atlas. The balance of digital entities, which comprise the Atlas itself, are made available to those consulting the CAA. (p. 15)

  This means that the Atlas will be used by those who consult it for their respective activities.

- The **originator** is the person to whom the Internet account issuing or the server holding the record belongs.

  There is no information found in the final report regarding whose or which server will be hosting the online Atlas. The description of the CAA project is presented on the University's Web site, but there is also information in the report stating that SCAR (Scientific Committee on Antarctic Research) is expected to take over the maintenance of the Atlas after the completion of the project.

5. Finally, a record must possess an identifiable context, defined as the framework in which the action in which the record participates takes place. The types of context include juridical-administrative, provenancial, procedural, documentary, and technological.

- The **juridical-administrative context** is the legal and organizational system in which the creating body belongs.

  The juridical-administrative context of the CAA is identifiable.

  The CAA Project is led by the Geomatics and Cartographic Research Centre (GCRC), an organized research unit (ORU) in the Department of Geography and Environmental Studies at Carleton University, Ottawa, Canada. The CAA is a key deliverable of a larger research project entitled Cybercartography and the New Economy (CANE), a 4-year project commenced in January, 2003. The CANE is funded by the Social Sciences and Humanities Research Council (SSHRC) of Canada under the Initiative on the New Economy (INE) program. A significant portion of the CAA Project infrastructure is funded by the Canada Foundation for Innovation (CFI).

  The research and development of the CAA is being carried out in partnership with a number of research laboratories at the Carleton University. The Project also collaborates with an international team of Antarctic scientists and multimedia visualization experts, who will share their expertise, laboratories, human resources and data with the Project.

  The CAA Project is subject to the rules and procedures governing SSHRC grant recipients. Also, the Project operates within the jurisdiction of the Carleton University and its rules and regulations.

  Regarding intellectual property and copyright, Antarctic Data fall under the Antarctic Treaty system and can be used, at no cost, for non-commercial research purposes.

- The **provenancial context** refers to the creating body, its mandate, structure and functions.

  The provenancial context of the CAA is identifiable.

  Information about the University can be found at
  http://www.carleton.ca/cu/aboutus/index.html.

  A link to its organizational chart is available at
  http://www.carleton.ca/cu/aboutus/organizational_chart/index.html.

  GCRC is under its Department of Geography and Environmental Studies.

  The mandate of the GCRC is to increase knowledge related to human interaction with geospatial information. The Centre fulfils its mandate largely through research in cybercartography and high resolution remote sensing for surveying and mapping and natural resource applications.

The GCRC's research focuses on Geographic Information Processing (GIP), Multimedia Cybercartography, Visualization and Remote Sensing, and the application of information and communications technologies in an international context. It has capacity in a broad range of activities in the GIP field in addition to its main research focus such as consulting expertise.

The CAA project operates on a task based approach, through a partnership of research laboratories at Carleton University, an international team of Antarctica scientists and multimedia visualization experts. In addition to Dr. D. R. Fraser Taylor, thirteen collaborators and an Advisory Board guide the project. The project also has a Project Manager and an Assistant Office Administrator.

See also the *Cybercartography and the New Economy Organizational Chart* on page 4 of the final report.

- The **procedural context** comprises the business procedure in the course of which the record is created.

  The procedural context is identifiable.

  Some of the procedures identified in the final report include:
  a. Incorporation of data from international sources
  b. Creation of content models using incorporated data
  c. Development of formal guidelines and procedures concerning the project
  d. Communications through online forum and meetings
  e. Development of projects architecture and framework
  f. Joint development of CAA user interface (UI)

  See also Figure 3, *Progress Diagram*, in the final report.

  Diplomatic analysis of **procedural phases** related to the Atlas activities under examination here can be broken down as follows:[13]

  a. **Initiative**: the introductory phase of any procedure is "constituted by those acts, written and/or oral, which start the mechanism of the procedure."[14]

     The initiative phase of the procedure of creating the Atlas started in 1999, when Dr. Vergani of Argentina first proposed such a project; it then formally adopted by SCAR-GGI in Sienna, 2001.[15]

[13] The phases of procedure as dictated by Diplomatic Analysis; see Luciana Duranti, *Diplomatics: New Uses for an Old Science* (Lanham, Maryland and London: The Scarecrow Press in association with the Society of American Archivists and the Association of Canadian Archivists, 1998), 115.
[14] Ibid.
[15] Pulsifer, P. L. and D.R.F. Taylor, "Project History," in *Cybercartographic Atlas of Antarctica*; Internet. Available at http://www.carleton.ca/gcrc/caap/presentations/pulsifer_taylor_jcadm7_07042003_arch_vers_ppt.pdf, accessed on January 26, 2006.

b. **Inquiry**: this preliminary phase "is constituted by the collection of the elements necessary to evaluate the situation."[16]

   The inquiry phase of the procedure of creating the Atlas is represented by Requirements Gathering in the Report, which includes, for instance, project team members' suggestions, user needs, technological considerations and consultations.[17]

c. **Consultation**: this phase is "constituted by the collection of opinions and advice after all the relevant data has been assembled."[18]

   The consultation phase of the procedure of creating the Atlas takes place in the process of designing the solution to answer the identified needs and requirements. It is an iterative approach where tools are evaluated for fitness and the design is refined. As the design evolves, it is compared to all stated requirements.[19]

d. **Deliberation**: this phase is "constituted by the final decision-making."[20]

   The deliberation phase of the procedure of creating the Atlas corresponds to the final decision made about the design of the Atlas. "It was decided to take an approach where the atlas modules would be written independently of the atlas' implementation details. Using this approach, the authors of atlas modules have a certain area of responsibility, while the atlas provider and atlas host have other clearly defined roles."[21]

e. **Deliberation control**: this phase is "constituted by the control exercised by a physical or juridical person different from the author of the document embodying the transaction, on the substance of the deliberations and/or on its forms."[22]

   There is no information found in the final report or all attached documents regarding the phase of deliberation control as defined by the above definition. In other words, there is no administrative control, i.e., no approve about the design from the principle investigator or some advisory committee responsible for directing the design and implementation of the Atlas. In the sense of scientific-technological control, however, the project's commitment to various standards and its employment of open sources can be viewed as the deliberation control towards the overall design decision made at the deliberation phase.

[16] Duranti, op. cit.
[17] Case Study 06 Final Report, Appendix R: Atlas Creative Process.
[18] Duranti, op. cit.
[19] Case Study 06 Final Report, Appendix R: Atlas Creative Process.
[20] Duranti, op. cit.
[21] Case Study 06 Final Report, Appendix R: Atlas Creative Process.
[22] Duranti, op. cit.

f. **Execution**: "the documents created in this phase are the originals of those embodying the transactions."[23] In other words, the execution phase results in the issuing of the first record capable of producing the consequences intended by its author.

   The Execution phase of the procedure of creating the Atlas is represented as the implementation of the design decision in the final report. One thing worth mentioning is that "the design and implementation phases are never completely dissociated since implementation details often affect the overall design. Often, early prototypes are assembled to prove or disprove some design assumptions. Also, review of available projects that could potentially be reused can often influence the design phase."[24]

- The **documentary context** is defined as the archival fonds to which a record belongs and its internal structure.

  There is no information in the final report regarding the recordkeeping practices of the creator. The CAA project in theory should form one sub-series within the GCRC series in the University fonds. The following information can be viewed about the recordkeeping practices within the CAA series.

  The digital components of the Atlas, mainly the content models and technical constructions (coding, web-deployment, etc.), are saved and maintained by content creators and technical specialists respectively. Within the project, however, they can be accessed through network and shared drive. The file management therefore is not standardized at the project level but subject to individuals' fashion of managing files. Nevertheless, it gives archival bond to each digital record, again, not to the online Atlas.

- **The technological context** is defined as the characteristics of the technological components of an electronic computing system in which records are created.

  The technological context is identifiable.

  The project will use the full extent of Internet technologies available up to date, and the Atlas will be deployed in a web-enabled environment using primarily open source software. Appendices H and I in the final report list all hardware and software being used in the CAA.

[23] Ibid., 116.
[24] Case Study 06 Final Report, Appendix R: Atlas Creative Process.

CONCLUSIONS

According to the above analysis, by satisfying requirement numbers 3 to 5, these documents—i.e., issue contents and programming codes—are identified as records.

Discussions on e-Records and e-Publications

The above analysis of assessing the Cybercartographic Atlas of Antarctica as a record demonstrates that the online Atlas only satisfy some of the conditions (e.g., the persons concurring in the process of its creation and some contexts) that are necessary for it to be considered as a record. The fact that its content (with fixed documentary form unable to be decided) is subject to continuous changing/updating, it does not participate in one action, and it does not possess an archival bond means that the Atlas only partially satisfies the definition of record. This diplomatic analysis thus concludes that the Atlas is not a record.

This analysis also reveals the characteristics of the CAA as a *publication*.[25] As stated in the above discussion regarding archive bond, the Atlas, upon completion, becomes a self-contained entity, standing on its own and does not require any other information to make it to be understood. It is presented to the public and publicly accessible, and it will have its own domain name, and a trademark with branding. Like any other publications, the messages it contains are self-explanatory and the meaning it conveys is complete. Even in a dynamic and interactive environment, every instantiation of the assembled data or every display responding to user inquiries is autonomous. Its publication status remains as long as it is hosted online and consulted as an Atlas.

[25] A "publication" is defined in the InterPARES 2 Terminology Database as: "Recorded information that is intended for communication and/or dissemination to the public at large."