A DIGITAL LIBRARY OF GAMBLING-RELATED PUBLICATIONS

By: Rhys Stevens, Alberta Gaming Research Institute Librarian
<rhys.stevens@uleth.ca>
Date: May 2005

INTRODUCTION TO THE PROJECT

Over the past several years, the Alberta Gaming Research Institute Library and the University of Calgary Image Centre have been collaborating on a project to digitize and archive a selection of historical and contemporary gambling-related publications. The project requires that digital copyright permissions be secured before items are digitized and made available online in their entirety through a digital library. The software used to run this digital library is called *DSpace* and it is hosted at the University of Calgary. As of May 2004, there were more than 239 items included in the collection. It can be accessed from the Institute's *DSpace* "community" at the following web address: <https://dspace.ucalgary.ca/handle/1880/79>.

Resources selected for inclusion in the collection are primarily comprised of monographs (books), technical reports, government publications, and graduate student dissertations. The nature of this particular digital library, however, also makes it possible to incorporate digital formats such as video & sound clips, data files, web pages, and presentation files (e.g. Microsoft PowerPoint). It is anticipated that this collection of materials will be of interest to academics and the general public; particularly those interested in tracing the growth and development of gambling in North America and elsewhere.

PROJECT GOALS

The two primary goals of the digitization initiative were:

- *To increase access to print & electronic gambling resources*, and;
- *To ensure long-term availability of print & electronic gambling resources*

To Increase Access to Print & Electronic Gambling Resources

The major aspects of the Institute Library's mandate are to support the research activities of the Institute and to function as a publicly accessible clearinghouse of gambling materials. The Institute's digitization initiative helps to fulfill these roles by making readily available a sizable selection of publications (particularly the "grey literature"[1]) with significance to the field of gambling studies. As no comparable collection of gambling-specific digitized materials is known to exist on the Internet, it is expected that this collection will play in central role in filling this identified gap.

---

[1] "Grey literature" can be thought of as publications from government, academics, business and industry sources which are not controlled by commercial publishers.

Many of the items selected for the collection were authored prior to the mid-nineties. As such, they were not originally made available on the Web and are generally not being included in collections of content being developed by commercial publishers (e.g. full-text databases of journal articles). This has meant that access to this literature is restricted to print and microfilmed copies available in libraries, archives, and private collections. To gain access to such items, individuals are required to either obtain them directly from such a collection or submit an inter-library loan request. Thus, the process of digitizing these materials holds great potential for improving this situation by using the Internet to widen accessibility.

Increasing the availability of gambling publications by including them on the Internet is, however, only one part of making them more accessible to researchers. It is also necessary to make provisions to aid in the most efficient location and retrieval of such documents. The *DSpace* software designed by MIT Libraries and the Hewlett-Packard Company helps with the accessibility issue as it is compatible with emerging metadata standards being used to facilitate document retrieval (e.g., the Open Archives Initiative (OAI) standards). Through the use of OAI-compatible standards, users are able to discover new resources by using specialized search engines which search across other academic digital libraries. One popular such search tools designed for OAI-compliant collection is called *OAIster* <http://oaister.umdl.umich.edu/> and was developed by the University of Michigan.

To Ensure Long-Term Availability of Print & Electronic Gambling Resources

Gambling scholars who incorporate materials from the Web (e.g., web pages, PDF files, Word Documents, etc.) into their research will often find that resources accessible online are relocated or removed altogether with alarming regularity. Koehler (2004) notes that it has been firmly established in the academic literature that Web pages are indeed ephemeral. A study conducted by Rumsey (2002) on web references cited in law journal articles found that only thirty per cent of them worked properly after a period of four years. Both authors indicate that, even though the Web contains an incredible wealth of information, it is not akin to a traditional archive since materials are not stored indefinitely.

The *DSpace* software used to create this digital library, however, makes it possible to store digital documents in a way that approximates a traditional archive. It intends to provide long-term access for the information contained in these files by making use of accepted preservation techniques for digital materials. One of the ways that this is accomplished is by using a "handle system" which assigns persistent identifiers to each item in the collection. Effectively, researchers who cite an item's "handle" can be assured that the item will always be available at that exact address. As an example, the handle <http://hdl.handle.net/1880/193> will permanently lead to the digitized version of the publication *Gambling in Canada: Golden Goose or Trojan Horse?* (1988).

MAKING USE OF THE DIGITAL LIBRARY

Accessing the Collection

The only requirement for use of the Institute's digital library is a high-speed connection to the Internet. Users can either connect directly to the Institute's *DSpace* community using the URL <https://dspace.ucalgary.ca/handle/1880/79> or first connect to the main *DSpace* site (URL: <https://dspace.ucalgary.ca/>) and then 'browse' to the 'Alberta Gaming Research Institute' community – *See Figures 1 & 2*.

Figure 1 – *Alberta Gaming Research Institute DSpace Community*

Figure 2 – *University of Calgary DSpace Home Page*

Item Metadata

In order to more effectively search the content of the collection, it is helpful to be aware of item metadata. Metadata is used to identify and describe materials in the digital collection. It can also document an item's function and use, its relationship to other information, and how it should be managed (Gilliland-Swetland, 2000). Each of the items in the gambling collection is assigned a metadata description. All such descriptions are then indexed by the *DSpace* software to facilitate efficient searching and retrieval.

Searching & Browsing the Collection

It is possible for users to search across all documents in the Institute's community of gambling materials by *keyword*. The keyword search searches both the metadata associated with an item as well as the digitized "full-text" of the item[2]. These types of searches are most useful when a user is attempting to discover resources when the specific document titles and authors are unknown.

Browsing is another possible way to locate documents in the collection. Users can choose to browse items by title, author, or published date (either chronologically or reverse-chronological). Such searches are particularly effective when particular information about a resource is known.

Matching documents titles are returned to digital library users who perform a keyword search or document browse. They can then 'click' on an item's title to reveal that document's unique web page which displays document metadata (i.e., title, author, publication date, keywords, abstract, etc.). A hyperlink to the file itself also appears on the page for users want to proceed to the actual digitized content. At present, the file format of choice for storing the majority of digitized content is Adobe PDF. There are, however, no restrictions on file types that can be included.

Special Features

For those interested in finding out when new items have been uploaded to the gambling collection, it is possible to "subscribe" to a reminder service that notifies users by email when this happens. Subscribing to the collection in this way ensures that an interested subscriber will be notified immediately when new documents are available.

Details on the process of selecting, digitizing, and uploading content to the collection are available in *Appendix A*.

---

[2] In order to capture the full-text from a digitized work, Optical Character Recognition (OCR) software is used to translate the images of letters and words into a format that a computer can recognize as text. Many, but not all documents in the collection have been OCR'ed.

EXAMPLES OF CONTENT INCLUDED IN THE COLLECTION

Though the majority of items available in the collection were produced in the twentieth and twenty-first centuries, there are a number of examples of **retrospective materials**. The earliest such document included is *By the King, a proclamation for the better regulating of lotteries within the kingdoms of Great Britain and Ireland* <http://hdl.handle.net/1880/444> which was a 1665 proclamation by King Charles II – *See Figure 3*. The earliest Canadian source appearing in

Figure 3 – *By the King, a proclamation*

the collection is the *Lottery for building a prison, for the Town and District of Montreal* <http://hdl.handle.net/1880/442> -- *See Figure 4*. This item is an image of early lottery ticket printed in 1784 to raise funds for construction of a Montreal prison. As these examples illustrate, making accessible digitized materials online has great potential for "unlocking" previously restricted access to a range of rare and unusual pieces of gambling history.
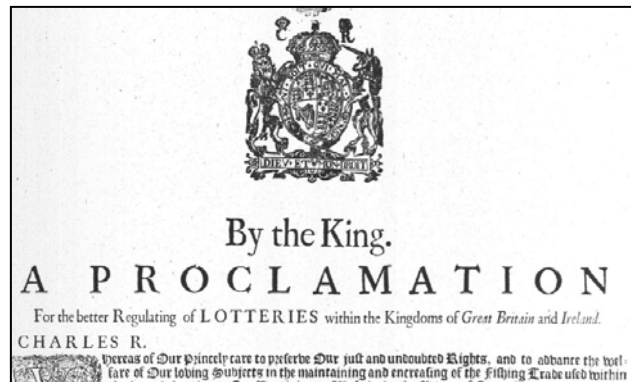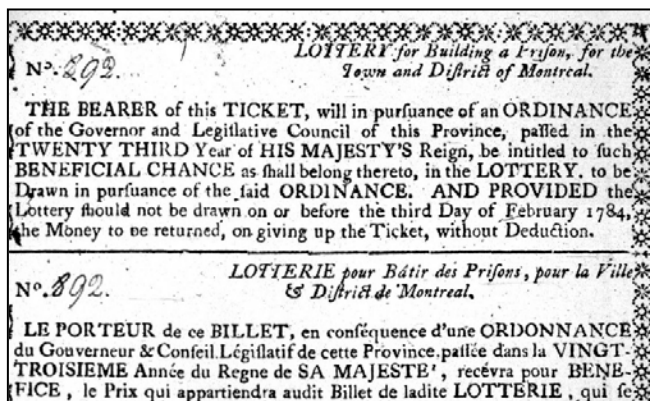
Figure 4 – *Lottery for building a prison*

While historical materials provide a rich source of materials for the collection, it should be noted that **contemporary items** are also excellent candidates for inclusion. Such items may never have been made available on the Web (e.g., *The history of the law of gaming in Canada* (1983) by R. G. Robinson for the Royal Canadian Mounted Police <http://hdl.handle.net/1880/1477>)

or may have once been available but have been subsequently removed (e.g., *Native gaming and gambling in Canada* (2001) by Kiedrowski & Associates Inc. <http://hdl.handle.net/1880/238>). Making such items permanently available can be invaluable for scholars and the interested lay public when attempting to comprehensively review the specialized gambling research literature.

CONCLUDING COMMENTS & FUTURE DIRECTIONS

The Institute's digitization initiative has, to date, proven that it has the potential to be an invaluable source of material on a variety of aspects related to gambling. As a large proportion of the historical publications related to gambling are inaccessible to many people, this particular digital library has great potential to increase access to these valuable materials. An additional benefit of the project has been to provide this access in a way that also facilitates stable and secure digital archiving of materials.

The nature of the *DSpace* software tool used to build the digital library also allows it to be other ways which can also provide substantial future value to both the gambling research community and general public. For instance, scholarly journal article "post-prints" can be deposited by authors and archived in the collection. This "self-archiving" is becoming increasingly popular in other scholarly communities and serves to increase the research impact of contributing authors. In fact, most scholarly journals already make provisions for such usage.

Although at the present time, it is difficult to accurately assess the value of this project to the research community. It is anticipated, however, that as the number of items in the collection increases and as awareness of the project grows, the lasting value of the collection will become readily apparent.

ACKNOWLEDGEMENTS

REFERENCE LIST

Gilliland-Swetland, A. J. (2000). *Introduction to metadata*. Getty Research Institute. Retrieved March 15, 2005 from http://www.getty.edu/research/conducting_research/standards/intrometadata/

Koehler, W. (2004) A longitudinal study of Web pages continued: A report after six years. *Information Research*, *9*(2) paper 174. Retrieved March 21, 2005 from http://informationr.net/ir/9-2/paper174.html

Rumsey, M. (2002). Runaway train: Problems of permanence, accessibility, and stability in the use of web sources in law review citations. *Law Library Journal, 94*, 27-39.

APPENDIX A - PROCESS OF SELECTING, DIGITIZING, & UPLOADING COLLECTION CONTENT

Step 1 -- Selection & Identification of Content for Inclusion

Selection of materials for inclusion in the digital library is based on a Collection Development Policy developed by the Institute Librarian in association with the University of Calgary Libraries - *See Appendix B - Site Standards and Editorial Policies*.

The Institute Librarian began the actual materials identification phase by locating citations to suitable materials from the following sources: records found in online library catalogues; gaming / gambling subject bibliographies, and; private collections of gambling researchers; existing online compilations of grey literature related to the subject (e.g. records included in the Gambling Literature Online Database <http://gaming.uleth.ca>). Four search iterations produced approximately 500 suitable items identified for inclusion. A *ProCite* citation management tool was used to create a database of citations collected. It should be pointed out that efforts were made to only include items that did not already exist in digital form on the Internet. Canadian items were also prioritized.

Step 2 -- Determination of Digital Copyright Owners

In stage two of the process, the copyright officer at University of Calgary Image Centre used the detailed citation listings developed in stage one to identify digital copyright holders for each item. Once contact information was located for each copyright owner, a request was made seeking permissions to produce and include a digitized copy of the item for use in the digital collection. Most owners were amenable to granting permissions for these items and were willing to provide them free of charge. Not surprisingly, owners of materials still "in print" did not grant permission or only permitted an introductory chapter to be included. In several instances, a small fee was requested by non-profit or charitable groups for the right to reproduce a digital copy.

In order to keep track of pertinent information (e.g. contact details, contact attempts, approval status, etc.) associated with every item identified for the collection, the Image Centre uses a specially-designed database. This database helps ensure that permissions are easily accessible if the legality of the digital copies is ever challenged.

Step 3 – Locating and Obtaining Physical Materials

Stage three of the digitization process involved Image Centre staff actually locating and obtaining physical representations of materials that had received copyright clearance for inclusion. This process was aided somewhat by the addition of library names and material call numbers to citation obtained during the identification phase. Fortunately, about twenty-five per cent of items selected were found to be available either in print or on microfiche from university library collections in Calgary and Lethbridge, Alberta. In

order to obtain materials not available from either of these two locations, the University of Calgary interlibrary loans services was utilized. In several instances, items were borrowed from the private collection of gambling researcher Dr. Garry Smith when unavailable elsewhere.

Step 4 -- Digitization of Materials

Once sourced materials were obtained by Image Centre staff, actual digitization of print and microfilm sources took place using high-quality scanning hardware unless items happened to already be available electronically. The process also included using OCR (Optical Character Recognition) software to recognize all document text appearing in materials.

This process resulted in the creation of numerous individual "page" files that were then be checked for quality. Once checked, each of the individual files was collated and packaged into a single larger file that represented the entire document.

Step 5 – Upload Digitized Files to DSpace Software

The fifth and final step in the digitization process was to make available the files created from the source materials. Files were uploaded to the *DSpace* repository by either the University of Calgary *DSpace* Manager or the Institute Librarian. During uploading, metadata was associated with each item (e.g., abstract, subject headings, etc.) in order to facilitate more effective search and retrieval. Once all required fields were added, materials became available for users of the digital library to use via the Internet.

The **Image Centre at the University of Calgary** is a multi-functional area specializing in the digital imaging of many types of original source material including print, microfilm or microfiche. Materials are scanned with a high definition scanner and the resulting images put through a rigorous quality assurance which checks for the clearest images possible. The images are then run through an Optical Character Recognition (OCR) process in which the images are read and recognized as text, making the document keyword searchable with a high degree of accuracy. An Adobe .PDF file can also be produced in order for documents to be downloaded and easily read. URL: Hhttp://www.ucalgary.ca/ImageCentre/digitization/digi.htmH

APPENDIX B - SITE STANDARDS AND EDITORIAL POLICIES

In order to lend authority to the contents of the digital collection, the following site standard and policies have been established.

1. To establish academic rigour and credibility of site content, **entire 'collections' of gambling materials will be digitized**. In other words, site content will not be 'selected' by project participants on the basis that it might be more important to website users.

2. **An Editorial Board will be established to address content issues and to identify digitization priorities**. The Board will include representatives from the three universities participating in the Alberta Gaming Research Institute consortium, as well as members from the Institute Board and Institute Council.

3. The Editorial Board will advise on the **order of items to be digitized**.

4. The website has been placed under the **imprint of the University of Calgary Press to clearly demonstrate the involvement of an Editorial Board** in selecting content and addressing content-related issues.

5. The Canadian Digital Gambling Resources site will be as **transparent as possible**. In other words, it will be made clear to clients using the website which organizations were involved in creating the site, and why content or collections were placed on the website.

    a. The concept of totality will be explained on the website, both in terms of assuring an academic researcher that no material has been purposely left off of the site and also to explain the importance of a historical collection.

    b. The only reasons that a known item will not be included on this site is that a copy cannot be located or because copyright permission to digitize cannot be obtained. In these cases, the title of the item will be included on the site with an explanatory note as to why it is missing.

    c. Copyright clearance is required for every item placed on the site with the exception of those in the public domain. Hand-signed copies of copyright clearance forms are preferred but digitally signed forms are also permitted.