THE STRATEGY OF THE NATIONAL DIGITAL LIBRARY OF LATVIA

Inga Grīnfelde, National Library of Latvia

Although the first digitization projects at the National Library of Latvia started 10 years ago, 2006 was the first time a designated budget of ~2 million Euros was allocated forming groundwork for the National Digital Library under the auspices of the National Library of Latvia.

Availability of funding in its turn enabled us to develop a consistent digital strategy for the next 6 years.

Mission

National Digital Library Letonica is a long-term government program whose mission is to extend knowledge about Latvian culture, history, geography, and economy thus promoting knowledge and understanding of the country in the broadest context to audiences involved in either academic or continuous life-long learning. We strive to establish ourselves as an integral part of the European and worldwide network of digital resources and involve public in its making.

Outcomes

The strategy seeks to achieve three long-term outcomes:

1. In collaboration with other memory institutions (libraries, museums, and archives) as well as education and research bodies a common Letonica information space or knowledge center will be established containing a continuously increasing number of Latvian information sources and accumulated knowledge in the digital format. The future generations, in turn, will have testimony of the intellectual, scientific, and creative accomplishments of our time.

2. Latvian population will have an easy access to Latvia's digital resources required for education, research, professional activities, leisure time, as well as exploration of one's cultural identity. In addition, the digital library's architecture will allow anyone to become a contributor thus turning from a consumer into an author.

3. As a component of the European projects, Latvian information sources will become part of broader European or global cultural heritage networks thus giving an opportunity to users throughout the world to get acquainted with Latvia, its history, culture, and people and foster Latvia's international recognition.

Target audience

Anyone interested in either academic or continuous life-long education.

Directions

All activities are divided into 5 major groups or directions:

1. Content
2. Systems architecture and ICT infrastructure
3. Services
4. Research and methodology

5. Public relations

Content

1. Digitization of the analog content
2. Preservation and indexing of digitally-born content

According to the mission statement, in the coming years the content will be limited to sources pertaining to Latvia, either geographically or by content. A closer look at documents falling into that category as well as the state of technologies allowed us to divide the whole volume roughly into two categories:

1. Digitization of the analog content

   1.1. Mass digitization

   Users expect no less than full text search in massive amounts of data. Mass digitization will entail large volumes of text documents that are self-describing and can be easily processed automatically, generally an OCR-ed text waiting for technologies that would enable better segmentation and markup.

   1.1.1. periodicals (in total ~7.000.000 pages of newspapers and ~6.000.000 pages of magazines)

   In July 2008 the first 300.000 pages of Latvian, German, Russian newspapers of 1930s-1950-s using a solution developed by Olive Software will be launched.

   1.1.2. books (~30.000.000 pages)

   Experiments using in-house recognition software are underway. Better display interface still needs to be developed.

   1.2. Targeted projects generally involving more elaborate and thorough processing, additional metadata fields, background research, keywords etc. for specific object types, e.g.:

   - rare books and manuscripts
   - images
   - sound recordings
   - selected musical scores
   - maps
   - digital repatriation
   - European (TEL+, EDL Local, Europeana) projects
   - educational content and digital collections for children

   By the end of 2008 there will be 34 of such smaller scale projects managed by both the National Library and its partners.

2. Preservation and indexing of digitally-born content

The law extends the commitment of the National Library as a national legal deposit holder to the digital environment as well. The same applies to archiving Latvia-related web content.

Besides the above functions, the library has also declared its intent to collaborate with major academic and research institutions to store a copy of their institutional repositories.

Decision regarding the so called *junk* content – specifically weblogs, is still pending.

Systems architecture and IT infrastructure

None of the above would be possible without proper infrastructure and systems that would allow to process, transfer, safely store, index, and present large amounts of data.

The backbone of any digital library is a digital object management system (DOM) custom-developed for the National Digital Library by Microsoft and its contractors and due in July 2008.

DOM is intended to:

- serve as a national digital registry

- be a trusted multi-million depository (including virus and integrity checking as well as easy format migration)

- provide a user friendly interface with Web 2.0 functionality, such as personal collections, commenting, ranking, RSS feeds etc.

- support de facto standards for data exchange protocols

- implement a digital rights management solution (DRM) on a file level

Its architecture is fully OAIS-compliant.

During the next 5 years we will also be working on:

- a solution for recognizing and indexing Latvian Fraktur (black letter typeface):
    - 1st stage – character by character recognition
    - 2nd stage – mapping to modern Latvian dictionaries for convenient searching

- Implementation of CAPTCHA or similar technology that would allow users to participate in the editing of inaccurately recognized text

- Resolution service for permanent identifiers

- **Resource consolidation or hybridization tool**

- Keyword assignment to images enabling more thorough indexing and searching

- Implementation of dictionaries (synonyms, old expressions) and controlled dictionaries (toponyms, personal names, abbreviations, etc.)

- Delivering to mobile appliances, such as e-book readers, smart phones etc.

Services

Digital environment allows us to extend traditional library services to much broader audiences and introduce a number of new ones, among them:

- Value-added content and full-text searching

- Digitization on-demand services

- 24/7 virtual reference

- Multilingual interface and metadata
- Automatic translation services
- Electronic payment for copyright protected work
- One-stop access

Research and Methodology

The leading position in the field has allowed the National Library to disseminate what used to be internal practices, workflows and guidelines to external audiences striving to put their assets online. This allowed us to ensure that most significant digitization projects follow the quality and interoperability standards that would enable them to be integrated with our, and further down the road – with European resource discovery tools.

For the last two years the library has been also offering a series of workshops and a set of guidelines to its partners. We continue to see our role as a center for best practice and methodology and intend to expand that role by offering a more systematic training that would be integrated into the academic program of the university library school as well as establishing distance education classes for partners in other parts of the country.

The other direction we will be working on is to enhance user satisfaction by improving searching experience and user interfaces. This will be done by carrying out an extensive research in semantic ontologies and search filters, natural language searching, indexing of audio-visual content, and user-centered design.

Large volume of accurately recognized text will enable researchers to develop tools for better language processing. The library is the principal contributor to the Latvian language corps.

A Research & Development section that was established in 2007, in collaboration with the University of Mainz has designed a computer linguistics course that will be offered to librarians and other interested individuals beginning Fall 2008.

Collaboration

No Latvian memory institution has a mandate to create an all-encompassing digital depository. For the benefit of the end user, this adds even more weight to our ability to find a common ground to ensure that digital libraries use platforms that are interoperable and data exchange standards that enable resource sharing.

As of Spring 2008, the National Digital Library has 16 partners throughout the country:

- two largest academic libraries
- three research institutions
- 11 public libraries
- one government agency

By 2009 we expect to have at least 20 partners that would allow us to continue to offer a diverse and engaging content.

We also intend to expand our cooperation with education institutions to provide targeted content in support of academic programs for different age groups.

Equally important is a public-private partnership (PPP). One of the ongoing projects together with a Latvian software manufacturer Tilde involves assigning geographic coordinates to images to be able to design applications that would place historic images on the maps.

Copyright

As everywhere else, copyright has been an issue that so far has put limitations to the digitized content. The latest amendment to the Latvian Copyright Law adopted in the beginning of 2008, stipulates free access to copyright-protected work within the library network and thus enables libraries to rapidly increase the digitization efforts.

Critical Success Factors

The success of the strategy depends on several factors:

- our aptitude to learn new technologies and form an ICT infrastructure to support the growing need for computing;
- ability to train and hire personnel with a specific set of knowledge and skills;
- ability to carry out organizational changes and allocate resources that would allow us to integrate digital formats and processes into standard library procedures;
- range of new products and services;
- strong digitization unit (including scanning and digitization of audiovisual formats) that would employ the best experts in the field;
- systems interoperability and common standards;
- long-term storage and access strategy;
- ability to collaborate with other Latvian memory institutions to develop common resources and methodology;
- consistent support by the management of the National Library and other memory institutions;
- reliable fiscal policy on behalf of the policy makers.

Strategy without action is a day-dream; action without strategy is a nightmare.

Old Japanese proverb