Research as a strategic tool for a research library

Antti Leino

Introduction

Libraries have been closely involved with higher learning and research since the dawn of scholarly literature, much longer than there have been universities. In Finland, what is now the National Library started out in 1640 as the library of the Royal Academy of Åbo, although both the university – nowadays named the University of Helsinki – and the library have their roots in the medieval ecclesiastical administration. In Finland as elsewhere, both technological and social innovations have over the centuries brought changes to the role of libraries. The most recent of these is the digital revolution.

The use of computational methods has been spreading slowly but steadily from hard sciences to other fields, and although the humanities have been a little behind they are catching up, so that today some universities have dedicated centres or even departments of digital humanities. At the same time, the general public has learned to expect services on the Internet. Both these developments require the library to change the way it operates. If nothing is done, or if the transition from paper-based to digital libraries is done badly, there is a clear danger that both university and general public libraries lose their user base. This can have serious effects for both the libraries themselves and for the society they serve.

While the National Library of Finland has for some time already had a nationally significant centre of expertise in its Centre for Preservation and Digitisation, it too has to keep up with the times. In the long run, professional expertise in digitisation is not enough: in order to keep at the top even nationally, one would need to aim at being at the top internationally. Furthermore, as a larger and larger amount of the existing material gets digitised the emphasis will necessarily change from the digitisation process itself to the later stages of making the material available for public. Fortunately, the National Library is not alone here. Some of the challenges are common to libraries and archives, while in others academic research can provide a basis for collaboration. As the National Library is a part of the University of Helsinki, a research-based approach has certain appeal.

Need for Research

The National Library of Finland serves the whole University of Helsinki as a research library and the entire nation as its national library. Still, the bulk of its scholarly users have come from the humanities, although here the term should be understood rather broadly to include for instance such fields as church or legal history, placed under the Faculties of Theology and Law.

The Faculty of Humanities at the University of Helsinki is in excellent shape, as seen from recent evaluations, although in general it has not been at the forefront of digital humanities. Still, some of the work done on computational and corpus-based research is world-class. Most notably in the present context, the research in language technology is highly regarded, and the university has been appointed as the national representing entity for CLARIN ERIC, the European research infrastructure consortium for language resources and technology. In a somewhat similar vein, the VARIENG research group is a major actor in corpus-based study into the development of the English language. Bringing the rest of the faculty up to this high standard is one of its strategic goals.

Looking at the digital world from another angle, computer science at the University of Helsinki has a long history of significant work in data mining, often in collaboration with other fields. In the humanities, such collaboration has ranged from stemmatology of mediaeval texts to spatial correlation of pre-Christian cult-related toponyms. There is some history of cross-disciplinary projects between humanities and other computationally inclined fields as well, such as Argeopop,

a joint research group between archaeology and population genetics.

Interesting and important as it is, most of the cross-disciplinary research is outside the scope of the National Library. From the library's point of view there are two major issues in the digital revolution. First, the publishing industry is moving towards digital formats, and second, people are getting more and more used to finding things on the net. These issues are closely connected to each other and to some extent also the changes that are happening in the research methods.

The demand for electronic access has led to more and more academic publications being available on the net. This development is not without its problems, and at times it has seemed that major scientific publishers are pricing digital versions of their journals in a way that hinders rather than furthers the dissemination of research results. On the other hand, small publishers may not have the resources for the transition from purely paper-based to digital publication.

Present-day scholars and general public have already come to expect digital services. They are used to finding information on the net, typically using search engines, to the extent that research articles available only on paper get cited noticeably less often than their digitally-available counterparts. Over the past decade or two, teaching new students to use libraries has become slowly but steadily harder, both because the students are less familiar with libraries to start with and because there are other, more familiar and thus more easy-to-use ways of finding relevant research.

The work is already under way in Finland to build a digital library infrastructure, and this project is not limited strictly to libraries: the FinELib consortium has as members also museums, specialised archives and research organisations. This makes sense, as after all many of the issues are common to libraries, archives and museums. The research organisations are in the consortium mainly because they collect corpora, but their expertise in using the data is also valuable.

The intimate relationship between academic research and research libraries means that librarians must have some familiarity with current methods used in the fields represented in the library, in addition to their main expertise in information studies and library science. A major methodological paradigm shift, as is happening now, means that especial care must be taken to keep up the contacts and collaboration between librarians and the research community.

In early 2011, the National Library spearheaded by its Centre for Preservation and Digitisation decided to tackle these issues with a dedicated cross-disciplinary research group. The group would ideally be headed by a professor in digital corpora and include research and support personnel from the library and organisations conducting research on both the humanities and digital technology.

Opportunities for Collaboration

A research-based approach to developing library services cannot be implemented by the library alone. A suitable set of high-class partners is required, and fortunately such partners can be found close by. In Helsinki, the Faculty of Humanities has a strong tradition in language technology. While the envisioned research group must have a strong cross-disciplinary slant, language technology is clearly one of the main fields involved, as the bulk of library material is textual in nature. Nevertheless, it must be kept in mind that the goal here is to serve all scholarly library users, not only those working in linguistic fields.

At the Faculty of Sciences, one of the strong research areas in computer science is data mining. Research in algorithmic data analysis has a long history, and it still continues to be at the top world-wide. Some of the research is conducted at the Department of Computer Science, some at the Helsinki Institute for Information Technology, an administratively distinct institute supervised jointly by the University of Helsinki and the Aalto University. Both institutes have a tradition of collaborating with other fields, to the point that some interdisciplinary programmes like bioinformatics have eventually matured into independent fields. Still, most of what the National

Library needs and proposes is application of rather than research into computer science.

The National Library itself already has a role in the digital world. It acts as the coordinator and central actor in the FinELib consortium, aimed at developing the National Electronic Library. A lot of the underlying work is already done, but there is still work to do before the digital collections are easily accessible to users. It has proved somewhat challenging to build the services to be offered to the general public, in part because the consortium consists of not only libraries but also other memory organisations, such as museums and specialised archives. The library also has a Centre for Preservation and Digitisation that specialises in not only the digitisation process itself but also issues related to storing and accessing digital material. Unlike the bulk of the library, the centre is in Mikkeli, a small provincial capital that has settled on a strategy heavily oriented towards digital services.

The City of Mikkeli has realised that the region has a significant cluster of expertise on information technology and worked diligently to support and build on it. The city's innovation and technology centre, Miktech Ltd, fosters the development of this expertise into economical growth, primarily in the form of Digitalmikkeli, a network of companies and other organisations involved in digital archiving and online services. Although research is not primarily interested in immediate commercial applications, the overall goals of Miktech are clearly in line with the plan outlined here, and their support is invaluable.

In forming a viable research collaboration, one of the main actors in the region is the Mikkeli University of Applied Sciences,[1] with its Master's Degree Programme in eServices and Digital Archiving. There is strong interest for research that would support the programme, especially if this would involve national and international research contacts and joint projects. In a shorter term, the programme has students – many with strong professional backgrounds – in need of research projects for their theses. All in all, there is a significant overlap between the expertise and needs of the Centre for Preservation and Digitisation and those of the Mikkeli University of Applied Sciences.

There is also other academic activity in the area: the University of Helsinki, Aalto University and Lappeenranta University of Technology each have a presence in the city. Their research units and teaching activities are loosely gathered into a slightly eclectic academic community under the Mikkeli University Consortium, whose administration is handled by the University of Helsinki. The Preservation and Digitisation Unit is also a part of this community, and this would make it easier for it to expand its activities. All in all, the city offers a promising environment for setting up research in the field of digital libraries and archives.

Joint Research

In early 2011, a decision was made at the Centre for Preservation and Digitisation to start working towards a professorship and joint research group involving interested parties in both Mikkeli and Helsinki. After initial contacts, a small working group consisting of personnel from the National Library and from the Departments of Computer Science and Modern Languages came to the conclusion that a more in-depth study of the viability was in order. The Mikkeli University Consortium was able to provide funding for such a study.

In the autumn of 2011, a questionnaire was sent to universities, memory organisations such as university libraries, museums and archives, and some commercial companies dealing with digital archiving and related online services. The responses led to a series of interviews and negotiations, so that by early 2012 it was possible to have a proposal concrete enough to apply for funding. Over the spring, this proposal has been polished further in order to have the professorship founded.

---

[1]The Finnish *ammattikorkeakoulu*, while usually translated as *University of Applied Sciences*, is not strictly speaking a full university but rather an institution of higher vocational education, much like the German *Fachhochschule*. It has the ability to grant bachelor's and master's degrees.

The current plan is to set up a research group, consisting of

- Personnel from the University of Helsinki / National Library
    - a professor in digital corpora, leading the group
    - 1–2 junior researchers, ideally working towards their doctorate or perhaps starting their post-doctoral work
    - a specialist in digitisation
- Personnel from the Mikkeli University of Applied Sciences
    - 1–2 senior teacher / researchers, with a doctoral degree and subsequent experience in teaching
    - students working towards their master's thesis

The task of this group will be to find new ways to refine raw digital data into more useful form and to develop methods for using the resulting corpora, either in research or as a part of making the culturally significant material available for lay users. What this means, in short, is to predict how people will use libraries ten years from now – and be part of the process of making it possible. The exact research plan will be formulated by the professor leading the group once she has been appointed, in collaboration with the project steering group that is projected to consist of representatives of the institutions involved.

The research group is going to be based in Mikkeli and supported by the city and the Mikkeli University Consortium. However, by normal university practice the professor will be employed by the Faculty of Humanities and placed within the Department of Modern Languages, the current home department of language technology. Accordingly, she will have duties in Helsinki as well. While commuting between the two cities is cumbersome, close contacts to the language technology staff and the rest of the faculty will be invaluable. It is also projected that the group will collaborate as necessary with other research groups, such as ones within the Helsinki Institute for Information Technology and the closely related departments at the University of Helsinki and Aalto University, and the School of Information Sciences at the University of Tampere.

The goal is to recruit the professor and the junior researchers so that the group can start working in early 2013, although university administrative procedures may impose a slight delay. Funding for the professor has been agreed on for the time period of 2013–16 and for the overall research group for 2013–14, although there are already some plans to extend this. The hope is that during the first few years the work will prove useful enough to justify a permanent research structure.