

DIGIBYB

**Digital Library Project
European Minority Languages**

Final report

Digibyb

Digital Library Project on
European Minority Languages

Final report

Fryske Akademy / Mercator – Education
Leeuwarden / May 2005

Contents

PREFACE 3

1 Introduction 4**2 Project organisation 5**

2.1 General 5

2.1.1 Aim of the project 5

2.1.2 Parties involved 5

2.2 The Digital Library Project in the Frisian context 6

2.2.1 Similar initiatives 6

2.2.2 Digital Library resources in Fryslân 7

2.2.3 Publishing houses and copyrights 7

2.3 The model 7

2.3.1 Function of parties 9

2.3.2 Functions of pilot parties 9

2.3.3 Future tasks and functions for the digital library 10

2.4 Justification 10

2.4.1 Technology and standards 10

2.4.2 Pilot content 13

3 Content of the websites 14

3.1 User groups 14

3.2 Objectives 14

3.3 Content plan 15

3.4 Quality criteria 17

3.5 Meta-data 18

3.6 Conversion 19

4 Time schedule 20

4.1 Project planning 20

4.1.1 Orientation phase 20

4.1.2 Implementation phase 21

4.1.3 Rounding-off phase 22

4.2 Project planning evaluation 23

4.3 Experiences and learning effects for the Fryske Akademy 24

4.4 Project continuation, follow-up 25

4.5 Budget 25

5 Technology 26**6 Summary 26****7 Bibliographical references 28**

Preface

The Digibyb project was subsidised by the 'Digitaliseringsfonds' of the KNAW: Koninklijke Nederlandse Akademie voor Wetenschappen (Royal Netherlands Academy of Arts and Sciences). Additional financial support was received from 'DENK' (Digitaal en Expertise Netwerk KNAW). Without these funds, this pilot project on a digital library of European minority languages would not have been possible. Therefore, we would like to express our gratitude for the wonderful opportunity we have been given. Our special thanks go to Jack Spaapen and Ella Uijtdeuwilgen of DENK for their support.

We would also like to thank the following people and institutions: our external advisor Bart van der Meij from Reekx for being a great help, especially in the initiating and starting stages of the project; Laurens Sesink and the team of i-Tor for their enthusiastic and specific technical support; Joris van Zundert for the brainstorming sessions; the members of the Advisory Board: prof.dr. Durk Gorter, dr. Willem Visser, drs. Cees Klapwijk, drs. Frank Bosmans, drs. Jelle Krol and drs. Menno Nonhebel; Tresoar (Frisian Historical and Literary Centre) and Bornmeer publishers; our colleagues of the Fryske Akademy and especially our former colleague dr. Alie van der Schaaf, who initiated the Digital Library Project.

We also owe a special word of thanks to Nieves Brisaboa of the Biblioteca Virtual Galega, who worked closely together with the team in the period prior to the III Mercator International Symposium to prepare a demonstration model for a European follow-up project. The model was presented during the symposium and proved to be a major success. The Biblioteca Virtual de Galega will be one of the partners in this future project.

Digibyb Project team
May 2005

1 Introduction

In 2003, we started our preparations to realise a Digital Library on European Minority Languages. The aim of the Digibybyb project was to carry out a pilot project on digital data storage and data retrieval concerning minority languages in Europe. The documentation of the project has been written in English with a view to the follow-up of this project, EMILL (European Minority Languages Library), and possible future European partners. The documentation consists of three separate reports:

- Final general report
- Technical Report
- Appendix

In a traditional context, libraries are usually relied on to gather specific information and background reading. In the past few years, however, the influence of Internet on information exchange and information gathering has been growing rapidly. Scientists, journalists, opinion makers, politicians, but also any private person can find an increasing amount of information on the Internet. The Internet does not give any warranties to ensure quality, but many users do not distinguish between sources - or simply do not have insight into the background of a source.

This is what makes it important for any serious institution to be visible on the Internet. For minority language communities, visibility on the Internet is far less evident. In an effort not to get lost in the information society, minority language communities should try to establish proper presentation and exposure on the Internet. This also holds good for the Frisian language community.

There are two sides to digital information concerning minority languages: information about the language and its position on the one hand, and information written in the minority language on the other. For the community itself, the latter is of great importance, but in order to be understood in the 'outside world', information about the language should be available in widely accessible languages.

Both types of information can be found in a well-established library: both types of information should also be present on the Internet in a digital library. Cooperation between minority language communities can be useful and is necessary for smaller players in the digital playground in these challenging surroundings.

2 Project organisation

2.1 General

2.1.1 Aim of the project

The aim of the Digibyb project was to carry out a pilot study on digital data storage and data retrieval concerning minority languages in Europe. This data gathering should result in a digital library, i.e. information containing sources and meta-data of sources, organised in such a way that it enables users to search, find and read the data. In other words: “Digital libraries are organisations that provide the resources, including the specialised staff, to select, structure, and offer intellectual access to interpret, distribute, preserve the integrity of, and ensure the persistence over time of collections of digital works so that they are readily and economically available for use by a defined community or set of communities.”¹

The contents of the data should concern minority groups in Europe. The resulting digital library should encourage information exchange between European minority groups and enforce the presence of these groups in the international digital community. The digital library is aimed at several user groups, such as scientists, students, teachers, journalists, policy makers, writers and interested laymen.

To realise the concept of a digital library, attention was directed towards the establishment of a digital library of Frisian, simultaneously enabling extension of the project towards other minority language libraries - or at least creating such a flexible concept that other digital libraries could easily be linked up with the Frisian pilot project.

2.1.2 Parties involved

The project was initiated by Mercator-Education, one of the three research and documentation centres in Europe and based at the Fryske Akademy, which studies education in relation to minority languages. Mercator-Education has a specialised library in the field of European minority languages. The project was financed by the Digitaliseringsfonds of the KNAW (Royal Netherlands Academy of Arts and Sciences). The project team that was subsequently formed consisted of staff members from the Fryske Akademy² in order to broaden the scope of the project and to gather expertise from several sources. Project team members were Alie van der Schaaf, succeeded by Cor van der Meer (project leader) at the beginning of 2004, Elly Albers (Mercator – Education library), Peter van der Meer (Fryske Akademy central library), Bob Boelhouver (Fryske Akademy ICT), Derk Drukker (Fryske Akademy ICT), Arjen Versloot (Linguistics Department of the Fryske Akademy). In the project’s initial stages, external advice was provided by Bart van der Meij from the library consulting bureau Reekx. This bureau has assisted in making the project’s goals and its planning operational.

Involvement of more specialists was organised in two ways: through the advisory board and through the so-called Information Studio. The composition of these two groups is described in appendices 7 and 8. Meetings of the Information Studio were organised in the early stages of the project (October 2003) and again in the final period during the III Mercator International Symposium (November 2004). The advisory board was consulted on a regular basis throughout the whole project period.

In the course of the project, practical collaboration could be established with “Tresoar”; the Frisian Historical and Literary Centre³, which resulted in access to a small subsection of their digital archives, i.e. digital images of old manuscripts. Furthermore, connections were established with publishing house Bornmeer, from whom we received digital PDF versions of recent book publications. For reasons of international cooperation and with the future project in mind, we collaborated with the digital library of Galicia (Spain) and linked up with their digital library BVG “Biblioteca Virtual de Galega”⁴.

2.2 The Digital Library Project in the Frisian context

2.2.1 Similar initiatives

The idea of disclosing library data in a digital form is not a new one. All around the world, libraries are working on the digitisation of their sources. Three approaches can be distinguished here:

1. scanning of existing materials (paper, parchment, etc.) as images;
2. collect primary digital data;
3. scanning and digitising (OCR) of primary non-digital data.

The first approach has the advantage of staying close to the original. Tresoar has published some mediaeval Frisian manuscripts on the Internet in this way. Still, especially hr-scans are not very Internet-friendly and scanned text cannot be searched.

The second approach is used more and more often, as more and more documents are processed in both hard and soft copies: both parallel and primary. The actual printed form can then no longer be considered to be the master copy. One example of this is formed by the digital newspaper archives.

The third approach aims at converting primary printed documents into digital form. This still is a very time-consuming business. When text interpretation is carried out automatically, it may be error-prone. In addition, it is difficult to get an exact copy of the original. A project using manual data entry in low-cost countries is the DBNL (Digitale Bibliotheek der Nederlandse Letteren¹). In this initiative, especially older Dutch literature is included (cf. section 2.2.3). Vast numbers of books have been digitised so far. The original's layout is lost, but page markers are included to enable a quick look-up. The website offers full-text search and searching by author, category, period or region.

2.2.2 Digital library resources in Fryslân

The largest library collection in Fryslân, Tresoar, is still in the exploring and starting phase of digital data disclosure. In fact, lots of documents have been scanned and / or copied on micro-films, to save the originals from damage through frequent use. The conditions for digital data retrieval through the Internet - at least the retrieval of source scans - are not particularly unfavourable, but little priority has been given so far to that approach.

The Linguistics Department of the Fryske Akademy has a large number of Frisian texts in digital form (Language Databases), processed according to principles similar to those used by the DBNL. These digital texts were created for linguistic research, and form by far the largest collection of Frisian texts in digital form. Given its primary goal, the collection contains only texts in the Frisian language and their selection is based on linguistic criteria.

When it comes to meta-data, the situation is a far happier one. Most institutes with library collections, including Tresoar and the Fryske Akademy, have digitised their catalogues,

Contemporaneous developments related to digital data collections, created by e-publishing, are very rare. We would like to mention private initiatives by Frisian poets and other activities of a totally different nature: the digital archives of the Frisian newspapers (mostly in Dutch). So far, the Fryske Akademy has done very little in this field: there are no digital Frisian scientific magazines, and larger documents are only published in the form of books, not as websites.

2.2.3 Publishing houses and copyrights

A striking problem in the creation of digital libraries is the question of copyright. Authors and publishers very often have commercial interests in the publication of texts: books, magazines and newspapers are likely to be sold. Publishing a text on the Internet has an effect on sales, which could be reduced or increased.

The complexity of the problem is further enhanced by the often double ownership of texts: both the author and the publisher may claim specific rights, and both have to agree on publishing texts on the Internet. For that reason, an initiative such as the DBNL⁶ concentrates on older literature, where the question of copyright is a non-issue. Especially in the case of DBNL, where the layout of the texts has been simplified, the DBNL project leader has noticed an increase in demand for the original, printed books following the publication of the text in the DBNL. Nevertheless, this argument is not appreciated in the same way by every publisher, so the problems of acquiring recent digital data for digital libraries remains problematic.

The following questions relate to technical aspects in this matter:

- to what extent is the digital version an exact copy of the original layout?
- can people retrieve the whole text at once?
- can the text be downloaded?
- is the text easily printable?
- is it easy to make a digital copy of the texts?

The more opportunities the library user is given, the more reluctant authors and publishers will become in releasing their digital texts.

In the practical case of the Digibyb project, we managed to overcome hesitations within our own institute. The project even acted as an inspiration to start formulating strategies and policies for e-publishing. We succeeded in establishing lasting contacts in this field with one of the main publishers of Frisian literature: Bornmeer Publishers.

Another important technical aspect is that of digital version management. Very often, the author does not possess the final digital version of a document, given the process of pre-press preparation and correction. In fact, the printer usually has the final version, and very often final printing files are not properly stored, as the printed version is considered to be the master copy (cf. section 2.4).

2.3 The model

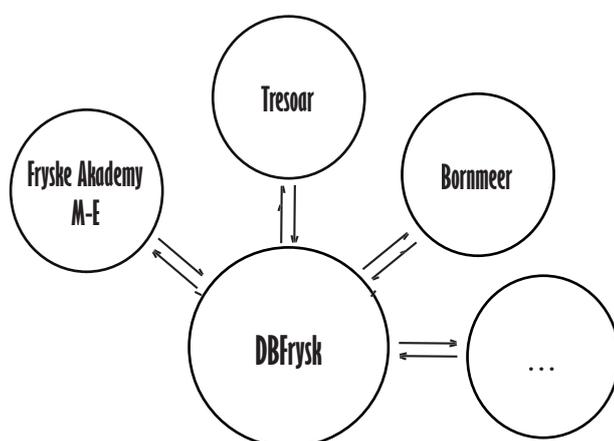


Fig. 1 Organisation model DBFrysk at www.dbfrysk.org.

The Digital Library Project has resulted in a two-tier organisation scheme. On the first level, the language-specific library has taken shape (cf. fig. 1). In the pilot, initiatives have been developed to make a start with a digital library of Frisian. This library addresses the regional user and will basically contain Frisian language data and data in Dutch. In the pilot project, emphasis has been put on developing the technical framework and on making a start with adding contents to the library. The results can be seen

The second tier is the European framework for a digital library on minority languages. The emphasis is on documents and data written in major European languages, and preferably in English. The European library itself should contain information about the languages rather than information in the languages involved. Furthermore, this European framework gives access to the various connected individual libraries (cf. fig. 2). The European framework can be found at www.emill.org.

The general approach adopted by both libraries is that of subsidiarity: information should be maintained and stored on the lowest possible level. By harvesting the data and creating library repositories, the information is collected on the European library level. The actual library data, such as texts, remain on the local servers of participating partners. The keyword here is open standard (OAI (Open Archive Initiative)); by providing the local data with meta-data in accordance with OAI standards, the library is able to detect the information and harvest it. (cf. Technical Report).

2.3.1 Function of parties

In both models, there is a need for central responsibility where maintenance of the library is concerned. The technical specifications in our pilot are described in the Technical Report. Of course, fine-tuning of these specifications with respect to the participants' actual situation is needed. The actual technical implementation can only be established on the basis of mutual agreement and discussion.

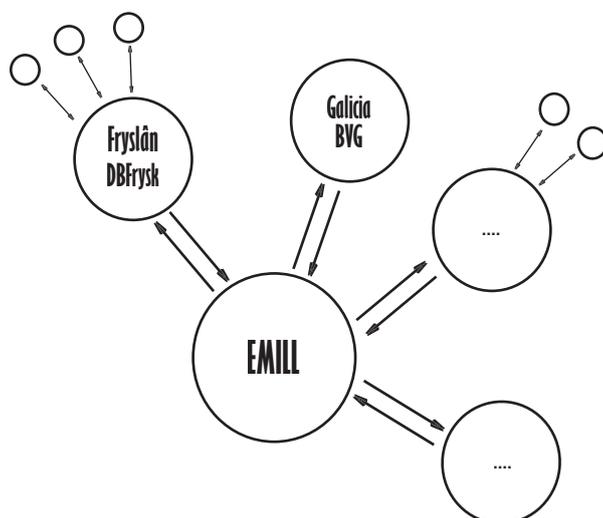


Fig. 1 Organisation model of the European Minority

The participating parties are responsible for making their own data available. This means that they have to enhance their primary and meta-data to be read by the central library server. This involves creating PDF versions of digital files, for example, and providing a mapping to a common meta-data format - in our case to Dublin Core. Once participants have prepared their data in accordance with the technical specifications, it is up to the central digital library service to ensure the harvesting of the data and to create search options and actual data retrieval in a user interface.

Besides technical task management, there is a need for content and quality management. Depending on the actual contents of the library (only digital books and articles, or perhaps also services such as e-publishing and community of practices) the management board will have to deal with the regulation of library use by its contributors and users. This includes setting up editorial boards for judging the content material offered, and safeguarding the actual goal of the library and its scientific integrity. Also, boards are needed to cover topics such as article reviews and administrative services for maintaining knowledge databases (such as databases containing contact information: addresses or telephone numbers of specialists in the fields). The subsidiarity principle requires tasks to be carried out at the lowest possible level. Once a partner is accepted in the project - based on general prerequisites and agreements - one may question the need and feasibility of controlling all of the contents offered by such a partner by a central board.

In a digital community, it is not even necessary to work with one single physical entity acting as 'headquarters'. Different tasks can be executed by different parties, such as hosting, harvesting, creating and implementing new search options, adding new data to the site, and so forth.

2.3.2 Function of the parties in the pilot

In the pilot, all functions were in fact carried out by the project group. As the actual pilot content was quite limited, there was no need to dress up complete editorial boards: in fact, very little was to be edited. Cooperation was sought with the library of Tresoar in Fryslân and with

publishing house Bornmeer⁷ in order to create at least a mini-model of the multi-party cooperation model aimed at (cf. Fig. 1).

On the European level, cooperation was sought and found with the Biblioteca Virtual Galega (BVG: the digital library of the Galician linguistic minority in Spain). This partner was willing to apply the standards for data retrieval as had been chosen and implemented by the project group (OAI, Dublin Core). This also enabled the project group to test technical integration with parties from different parts of Europe. As the Digital Library project is meant to be a growing model for European minority groups, it was very important that this aspect could indeed be implemented successfully.

Furthermore, to avoid complications regarding digitisation of printed material and the question of copyright, the content was limited to in-house Fryske Akademy material. In fact, effort was put into developing technical conditions for the digital library, such as creating OAI-repositories from existing databases by conversion of locally stored material of the library software programme. So, by defining the requirements for the technical standards, the project group also defined the contents of the library through exclusion of non-feasible goals.

2.3.3 Tasks and functions for the digital library in future

Two different levels should be distinguished. The European level (European Minority Languages Library : EMILL) is explicitly involved with European minority groups and concerned with increasing information exchange between those groups. In this field, a continued involvement of the Mercator centres⁸, and Mercator-Education⁹ at the Fryske Akademy in particular, is intended. As the technical prerequisites have been realised at the end of the pilot, the time has now come to extend the library services and to find more participants and contributors to broaden the contents of the library.

For the regional Frisian part, the future role of the Fryske Akademy should be reduced. It will be a continued task to convert material from the Fryske Akademy in accordance with digital library standards, but the Fryske Akademy cannot take a central position in maintaining the technical services for a digital library on a provincial scale. In fact, this task would be much more appropriate for the existing library service at Tresoar. It is Tresoar that keeps the largest collection of printed material on Fryslân. This is why we have been establishing contacts and have tried to direct their interest towards the possibilities and opportunities of digital library services.

2.4. Justification

2.4.1 Technology and standards

In this section we will account for the choices made regarding techniques and technical standards employed. The reader is referred to Chapter 5 and Appendix 3 for more technical information and specifications.

A digital library and OAI

The project group chose to create a digital library to be run like the model presented in the

previous section. This means that we wanted to have a central library with several parties, who keep and maintain their own data. The alternative would be to gather all the material ourselves and store it on a central data server. This is in fact the option that has been chosen by the Dutch DBNL. As the goals of the DBNL are different, namely creating a record of Dutch literature, this approach would cause several difficulties in the context of a multi-party model. Not only does the DBNL model require extensive server capacity, but also the problem of version management becomes an issue when one works with several information sources. This holds especially good for information that needs to be updated often or at least regularly: when and how will those updates be copied to the central server?

In order to create a dynamic digital library, a special tool is needed that enables a user to read remote data, retrieve and index them - so-called data harvesting. To make this possible, primary data needs to be translated into some kind of open exchange format.

In a multi-dimensional project where several partners and participants need to work together in a network environment, it is absolutely necessary that all agree and comply to the same formats. In order to facilitate the decision making process, these formats need to be widely accepted standards. Most of the time, it takes a lot of work and effort to convert the information to these standards. Some European projects have experienced that if these formats are unique formats (and therefore not true standards), participants often show a lack of willingness and effort. For this project, we have primarily opted for open standards, such as OAI (Open Archive Initiative), XML (eXtensible Markup Language), HTML (Hyper Text Markup Language) and DC (Dublin Core). Widely accepted and applied formats such as PDF (Adobe) could also be used.

We have chosen OAI (Open Archive Initiative) because of its concept, because it is an open standard and because it is rather widely used. So, the data and structures of participating libraries and data collections will be available through the OAI protocol. Consequently, the OAI data can be read by the harvesting software. For any extension of the project, the use of OAI is compulsory.

In addition, we wanted to ensure the possibility of creating repositories of libraries, of extensive search options, the possibility of including multi-media data (sound, images and movies), user registration and user tracking, and also some other services such as digital publishing and community of practices. Such a software package can only function when the parties involved are willing to prepare their data for data harvesting. This means that they have to convert their local formats into formats that can be used by the library by means of some general exchange format. This applies to all kinds of data: primary data, meta-data, and additional information.

On the other hand, the programme to be used should preferably be able to handle a wide range of data formats. This implies that participating partners have considerable freedom to choose their own data format- also within the given application.

Application choice

In order to choose an application that complies with our demands, we drew up a list of demands, ranked according to three criteria: 'required', 'desirable', 'not-essential'. Several applications were picked and checked for suitability. A full record of this selection operation is found

in Appendix 3. Finally, i-Tor was selected¹⁰. There were several reasons for this, some of which were quite obvious:

- i-Tor complied with most of the demands,
- it makes use of the standards mentioned before,
- it was developed by a KNAW (Royal Netherlands Academy of Arts and Sciences) partner,
- it was the only application developed in the Netherlands, something which allows for regular personal contacts,
- it was very budget-friendly; i-Tor was in fact interested in cooperating with our Digital Library Project to improve and promote their product.

In the dynamic world of digitisation, very few things seem to be designed to last forever. So, the continuity of i-Tor is not self-evident. If the package ceased to be maintained and developed in the course of time, the project would face serious set-backs, but on the other hand the core itself would not be involved. Databases with primary and secondary data will not be affected if i-Tor were to be discontinued. Any other open source software programme offering the possibility of harvesting OAI repositories will be able to take over most of the functions. Of course, specific i-Tor interfaces such as 'search screens', for example, would be lost.

On the other hand, for future projects, the option of applying other software packages remains. The Biblioteca Virtual Galega, for example, uses a different open source software package, which also looks quite good. This package is functionally similar to i-Tor, so using that - or any other package - is still an option for the future.

Other choices

One of the central data categories in a library is that of meta-data: data about primary data, which makes it possible to search and find data. Meta-data are normally stored in catalogues.

Nowadays, most library catalogues have been digitised. The characteristics of categories applied in the catalogues may be quite diverse, although information concerning authors and titles is very common. Still, even here, the definitions may be quite different in cases with less obvious authors or titles. In order to make the different data sets united in the Digital library searchable, there is a need for an international standard for meta-data.

We have opted for Dublin Core, because of its wide applicability and usage. This implies that local catalogue categories have to be mapped to Dublin Core. To ensure easy and clear access to the information of the participants, this standard is almost compulsory. The wide usage of Dublin Core is an advantage and will probably accommodate this aspect smoothly. (c.f. section 3.5)

For data representation from digital book collections, we had several possibilities: HTML, XML-TEI, RTF, Worddocs and PDF. In fact, none of the data sources had been put in that format. It proved to be very laborious indeed to convert other data formats such as ASCII, ANSII or Word-Perfect into HTML or XML files with relevant and appropriate layout. Therefore, we chose to convert our data to PDF format.

This format has the advantage of being quite consistent in its presentation: books can easily be presented in their original layout, which is important for digital copies of primary printed texts.

The creation of an exact PDF copy from a digital file still often proved to be laborious. In many cases, the latest digital version stored by the author is not exactly the one that was used for final printing. Practical experiences obtained in that process are described in section 3.6.

2.4.2 Pilot contents

The intention of the pilot was not only to create a technical environment for a digital library, but also to organise a try-out for content management. For that reason, we defined several selection criteria and quality standards for the contents, which should be:

- correct and unoffending,
- relevant for the purpose of the project,
- not commercial and freely available,
- recent and up-to-date,
- properly laid out,
- well-documented (with regard to meta-data).

Some of these criteria have a general character and deal with the library function. This would require a central editing board. The project group was intended to have this function in the pilot stage. To ensure the relevance of the contents, we developed the idea of setting up boards of specialists who could define long lists and short lists of documents to be included in the library. As the Fryske Akademy is a scientific institute dealing with topics such as minority languages and linguistics, it was the most obvious option in the pilot to look for specialists from within the organisation. This was also considered to be a desirable long-term construction (cf. the organisation model). For several other topics, specialists would have to be found from outside the Fryske Akademy.

In practice, it turned out to be too ambitious to establish such a content management infrastructure during the pilot stage. It proved that the technical constraints were much narrower than the qualitative ones. To avoid any copyright problems, which tend to be very time-consuming, we chose to use mainly Fryske Akademy material. Even there, we had discussions about the availability of recent materials.

Our source material could be divided into two main groups: primary Frisian texts (literature) from earlier centuries (free of copyright and already digitally available) and scientific publications written by the Fryske Akademy staff. On the basis of content criteria (relevance etc.), we defined two short lists from these two sets of material. The actual implementation of these two lists and transforming them into library content proved to be difficult and time-consuming, so that the rest of the activities became primarily determined by the technical constraints and the availability of files. The reader is referred to Appendix 4 for the description of the actual pilot contents.

3 Content of the websites

3.1 Usergroups

The Digital Library Project deals with two sub-projects:

- DBFRysk: a digital library of the Frisian language and Fryslân,
- EMILL: European Minority Language Library, dealing with items of general interest.

By consequence, the project's user groups vary. DBFRysk will primarily be directed towards the users' own language communities. The Frisian language is the immediate object - and often the vehicle - of the digital library. Its final aim is to offer users the Tresoar collection concerning Fryslân (in fact the major part of the collection) in a digital form. Thus, the user group is also one and the same: anyone interested in Fryslân, the Frisian language and in any scientific activities dealing with it. One could think of scientists (Tresoar is in fact taking the role of university library for the Fryske Akademy), students, journalists, policy makers, and any interested laymen.

For EMILL, user groups can be defined in similar terms, but the project's scope is quite different. People from other countries interested in Frisian matters will - in most cases - hardly be helped by Frisian literature texts in the Frisian language. What they are interested in are descriptions of Fryslân, its society its language and so on in a broadly accessible language such as English for example. The group of users interested in such specific information per country is much more limited than the potential user groups for, say, DBFRysk, but given its international approach it may still reach a substantial number of people.

Given the limited size of the pilot, especially regarding the amount of content incorporated at this stage, we can hardly talk about any substantial user group at the moment. When we made the contents available on the respective sites, we did so with the desired user groups in mind.

3.2 Objectives

In the introduction, we briefly defined the aim of the project as: "[...] to carry out a pilot project on digital data storage and data retrieval concerning minority languages in Europe. This data gathering should result in a digital library, i.e. a body of information containing sources and meta-data of sources, organised in a way that enables searching, finding and reading the data, [...]". Following the subdivision between DBFRysk en EMILL, this general aim can be specified in greater detail.

In an ideal form, DBFRysk as such mainly represents a digital version of Tresoar, a historical and cultural centre of the language community. The main constraint lies in the physical task of making all the printed material digitally readable in any form. The digital software package used offers an additional possibility of providing other functions, such as digital publishing by peer reviewing and community of practice, also for the Frisian user group as such.

For EMILL, there is no such existing library like Tresoar that lends itself for digitisation. Of course, in the Mercator network documentation centres¹¹, a lot can be found that would be suitable to be included in the EMILL library. However, the aim of EMILL is to establish a more active role in stimulating information exchange between experts within and among the several European minority groups. The social and political topics touched upon in these regions show a great deal of resemblance, and therefore quick information retrieval concerning those subjects is of supra-regional interest.

As information exchange is a central item in the aim of EMILL, additional functions such as databases with up-to-date lists of specialised organisations and people, or thematically defined user groups are of even greater interest in this part of the project. Although this aim was maintained during the pilot and although technical requirements for this kind of functionality were met, only limited examples of information and services of this kind could actually be implemented in the pilot stage.

3.3 Contentplan

On the basis of the aims of projected user groups, a content plan (CP) was written for the digital library. The CP is added to the report as Appendix 2. The CP can be seen as a guide for the Fryske Akademy part of the project, and contains useful tips for any other party willing to participate in the Digital Library project.

The thematic fields considered to be essential for the project include linguistics, literature, sociolinguistics, history and culture, education, media, language policy and legislation. These are also the fields in which Tresoar plays an active role when it comes to gathering analogue data on Fryslân. Following the differences between DBFrysk and EMILL, the meta-language of EMILL is English, and texts should preferably be accessible in that language - or any other 'major' European language, such as French or German. It may also be clear that Frisian literature is of little interest to most of the EMILL users, as a majority of them will probably not speak Frisian.

Specific attention should be given to specialised data for international use, such as up-to-date lists of institutions and specialists in specific fields concerning minority languages, or addresses of European institutions dealing with language policy and financing projects. This requires a dynamic editorial staff able to keep addresses, including Internet addresses and links, up-to-date.

Intended contents of the pilot

It cannot be the aim of a pilot to realise all final goals. For that reason, we wanted to try out several types of content in order to:

- experience the technical constraints and demands;
- find out the workflow involved in obtaining the actual different types of content.

Other limitations included the availability of data and the problem of copyright concerning a lot of recent data.

The Fryske Akademy has a huge digital data collection of Frisian literature. This collection offers nearly 100% coverage for Frisian literature before 1800 and includes a substantial collection of post-1800 texts. The collection was built for linguistic purposes and set up in a period that offered far fewer technical facilities than is the case today. This means that little attention was paid to purely bibliographical aspects such as original layout. Also, only a limited amount of meta-data was collected (compared to today's library standards). For the modern data, we were confronted with copyright constraints, but it was seen as a challenge to try and retrieve these data, stored in linguistic databases, for the purpose of setting up a digital library. It was considered both a technical challenge and an opportunity of offering at least some form of Frisian language texts.

Other content material was expected to come from our own institute in the form of scientific publications (cf. section 2.4.2.). These include English-language work on the Frisian language and on sociolinguistics. They could be considered good examples of data relevant to EMILL. Furthermore, a list was included of European organisations working in the field of minority languages (ORGA database), a typical EMILL service product. This database has just been updated.

Another interesting content object is formed by the Regional Dossiers written by Mercator-Education. These dossiers are written in English and focus on the situation of education in the respective minority language communities. These also form a content product typical of EMILL.

It was intended to seek cooperation outside the Fryske Akademy with other digital collections, in Fryslân as well as in other parts of Europe. As certain difficulties were expected in this field, no firm goals were set in that respect.

Contents realised

The language database with historical Frisian texts was successfully connected to DBFrysk. The selection of suitable texts and the conversion of the meta-data appeared to be somewhat difficult, but in general this operation was successful.

The conversion of scientific publications turned out to be a problem of a mainly technical nature. Version management, layout, outdated digital formats, graphics, and page numbering were problem areas. As a result, only a limited collection could be set up. However, a substantial addition to the collection could be realised in the final stage of the project by extra staff investments. We also succeeded in finding one external publisher willing to provide modern Frisian literary texts in PDF format to be included on the DBFrysk website. This editor is particularly interested in the role DBFrysk could play as a digital archive for contemporary Frisian books.

The Mercator-Education Regional Dossiers could easily be converted into PDF and added to DBFrysk and EMILL. Also, the incorporation of the ORGA database (database of organisations) was rather easy - from a technical perspective, that is. Nevertheless, maintenance itself of the contents still needs to be incorporated in the workflow.

The pilot was relatively successful in testing and elaborating cooperation with external partners. Inside Fryslân, we managed to establish a technical link with a - be it small - digital subcollection of Tresoar, which can easily serve as a growing model. Looking elsewhere in Europe, we

found the rich digital library Biblioteca Virtual Galega ready to apply the techniques for shared information exchange in EMILL to their data collection. This may also serve as a growing model and example for the potential of EMILL.

Future content extension

The future extension of both Digital Library projects can be defined in three fields:

In paragraph 2.3.3, it was already described that Tresoar would in fact be the perfect party to fill a digital library DBFrysk with lectures and literature on Fryslân and the Frisian language. Of course, wherever the Fryske Akademy has digital texts available (for example post-1800 literature in the language database) it will participate and support such developments. For practical purposes, formulating a road map would be desirable towards full digitisation. The map should be based upon needs as well as on availability of material and also on digitisation resources.

To broaden the scope of EMILL, cooperation should be sought with other minority language communities in Europe. In the pilot, extensive experience has been gathered and a basic technical framework has been created which can be of great interest to other parties. Moreover, the Fryske Akademy is willing to extend the project further to accommodate other current initiatives.

In order to substantiate the international part of EMILL, further steps have to be taken. These include active data gathering concerning experts and institutes that would be interested in participating in information and knowledge exchange. It is precisely this field that the Mercator network of research and documentation centres is active in, so a further development in that direction would fit Mercator's goals and ambitions perfectly well.

To be a useful tool for the various user groups aimed at, a future extension of the contents is absolutely essential. The present libraries only have a basic content and mainly serve as examples.

3.4 Quality criteria

Besides the demand for an increase in content - if people do not find anything interesting, they will not consult the library again - another essential aspect of a library is its quality. If you find too much irrelevant material, you will avoid that site in future.

For that reason, it is necessary to safeguard and maintain the quality of the library. Some criteria for quality policies and their organisation have already been described in section 2.4.2. It is important to stress once again that in a cooperation model without a strong, well-equipped and well-staffed central office it will be difficult to control every quality aspect related to the various digital library projects. We would like to emphasise the subsidiarity of responsibilities: each partner is responsible for its own part, its own task and its own information. Partners as a whole can be judged according to mutual agreements. The CP offers a set of criteria to take into account in the definition of quality criteria.

Layout quality aspects

There are two technical aspects concerning text layout we were confronted with during our pilot. We shall briefly discuss them here.

First, there was the general quality of text layout. In today's visual and graphic society, it is very important that both the site and the texts visible look professional, and that they are easy to read. Although this is partly a matter of style, some general rules concerning readability have to be respected here.

A second aspect concerned the relationship between the text layout in the digital library and the printed publications. This is especially important when we deal with work that has primarily been published in printed form. From a bibliographical point of view, it would be desirable to have an exact match between the two, so that references made to pages in a publication are valid for both publication versions.

In practice, we found ourselves confronted with several problems when we tried to realise such a fixed relationship (cf section 2.4.1.). In many cases, such a relationship can only be realised when one either invests a lot of time in a digital 'reconstruction' of the original, or simply uses scans from the original (cf section 2.2.1). When it comes to publications with several illustrations, a complete digital reconstruction is too labour-intensive, and it will be too expensive to ensure large-scale content generation in the near future. A combination of the two approaches will probably prove to be the most flexible for future activities.

3.5 Meta-data

Mapping to Dublin Core (un)qualified

A distinction should be made between Dublin Core qualified and Dublin Core unqualified. Besides the so-called resource synopsis, the Dublin Core term for what has been described, three other synopses are of importance in Dublin Core terminology: 'element', 'qualifier' and 'scheme'. A resource is described with the assistance of a maximum fifteen different elements, for instance the element 'creator' which describes author, photographer etc. Every element can be deleted, but also show, if necessary. The order is not important.

An important rule of thumb in adopting Dublin Core is the so-called dumbing down principle: deleting a more specific indication of information inside an element must never lead to loss of information on a higher level. This is the reason why Dublin Core qualified still is a controversial extension. Dublin Core qualified has an extension to the 15 basic elements defined: the so-called qualifiers. Time and place can be specified by the elements 'date' and 'coverage'. Finally, an element can be completed by using a scheme such as that of the Library of Congress or the Dutch Basic Classification Scheme. By using the element 'subject', it can be indicated that a keyword from a thesaurus is being used and not just any arbitrary word. In the element 'scheme', the choice of thesaurus will be laid down.

Mapping to Dublin Core in the DigibyB project

To convert the existing meta-data systems in the Fryske Akademy (BRS-TTSearch, MYSQL and ACCESS) to Dublin Core, it is necessary to define which field matches with a corresponding

element in Dublin Core (completed with a qualifier and/or a scheme). This process is called mapping. For the conversion, we used "Best Practices" developed by CIMI, BIBLINK and other organisations. In the end, the information is exported to the OAI repositories.

Experiences

To date, hardly any experiences are known of users who have searched the records mapped to Dublin Core. The existing 15 elements of Dublin Core were expanded with DC-qualifiers, which is why almost all fields could be mapped.

Unfortunately, loss of some information is unavoidable in the process of mapping. The following list explains why this should be so.

- Not all present information could be admitted, but the most important information was defined.
- Limited use of Dublin Core is possible by describing resources such as books, paintings etcetera
- Different fields of the Bibliographical database of Mercator-Education were presented one behind the other in the same Dublin Core field. The field 'Language publication is about' and the field 'descriptor' are both mapped to the field 'subject' of Dublin Core. Specific information can get lost for that reason.
- By mapping the Database of organisations with name, address and contact information we decided to realise a mapping with Dublin Core elements while in fact Dublin Core is not designed for people and organizations.

A thorough knowledge of the fields and databases is absolutely necessary for the mapping process. Dublin Core only has a limited number of fields, not all information can be converted. Dublin Core is easy to use and implement, it has become a popular standard, it is in widespread use among many organizations. Dublin Core particularly useful for multi-discipline meta-data publication and dissemination (for mapping examples see Appendix 5).

3.6 Conversion

As base a Cd-rom is used with files of the various publications written by Fryske Akademy staff members. These files are stored in different formats - for instance WordPerfect 4.2 or Postscript. Several programmes were used to convert the files. Within the framework of this pilot project, attempts were made to make the layout of the files resemble the original printed version as closely as possible. Experience grew by calculating time during the conversion process of the older material and duplication of the original printed version.

Error handling

During the conversion process, it appeared that printed publications are often changed after the final digital file has been made available. Errors are quite often discovered just before the printing process, and these errors are often fixed at the printing office. This is one of the reasons why final digital versions of the publications differ from the actual printed publications. It is mainly the layout that has been changed in such cases, which means that the work has to be redone. Most frequent problem areas include page numbering and parts of pages which are missing, as is often the case with maps and pictures, for example.

Time schedule for file conversion

Conversion and error checking take a considerable amount of time. Time investment depends on the number of pages of the book and the number of errors in the file. The average amount of time for a complete conversion and error check is approximately 2-6 hours. Sometimes, it proved necessary to double-check a file when many errors were discovered.

4 Time schedule

4.1 Project planning

The project was divided into three functional parts: an orientation phase, an implementation phase and a rounding-off phase, originally in time scheduled as follows:

Orientation phase	September 2003 – January 2004
Implementation phase	January 2004 – July 2004
Rounding-off phase	July 2004 – October 2004

In the summer of 2004 it became clear that the project could not be finished in good order before the deadline of October 2004. We suffered quite some delay in the implementation of the software (i-Tor), mostly caused by technical hindrances such as a change of project leader and staff holidays. Luckily, an extension request was quickly approved by the KNAW, with the new deadline for the project set to 30 April 2005. This could be realised without raising project funds.

4.1.1 Orientation phase

The project started with so-called Starting Notes (2003) (Appendix 1): a starting paper summarising the goals of the project, user groups, general technical layout and possible partners for cooperation. This paper was discussed with specialists such as librarians and publishers in a special session called “Kennisatelier” (Information Studio). Participants of that session had the opportunity to reflect upon the concept presented. A questionnaire concerning the aim of the digital library was filled in by all visitors (Appendix 8). Although an impressive number of participants was present, the number and the actual contents of the reflections that were returned proved to be quite limited. In fact, none of the participants from outside the Fryske Akademy made any substantial contribution to the project, apart from the publisher Bornmeer and the library of Tresoar, in both cases after renewed contacts.

At the end of this stage, the project group had some idea about what to do and how to go about it. In the meantime, the Fryske Akademy directors had given their approval to making own publications – with a minimum age of 2 years - available for publication in the digital library.

Content Plan (CP)

Based on the general planning stages from the Starting Notes, a detailed content plan was written (CP, Appendix 2). This included considerations on a number of topics: what should be included in the library; according to which quality standards and criteria should this be done; how to organise content management; how to manage the publications' meta-data, and finally the organisation of other functions of a digital library - such as information services, exchange and e-working.

Writing the Content Plan was a very useful exercise in thinking about all kinds of aspects related to a digital library. In the actual pilot practice, it turned out to be slightly too ambitious. Most of the careful content planning described in the CP could not be effectuated in practice during the pilot project, but will undoubtedly prove to be helpful and effective in the longer term.

Technical considerations

For the digital library, a framework had to be found that would realise all the ambitions laid down in the Starting Notes and in the CP. From those two documents, a list of criteria – ranging from 'essential' to 'desirable' - could be formulated for a future software programme. The considerations are laid down in the so-called "Applicatiekeuze" or "Application Choice" (see Appendix 3).

On the basis of this list of demands, several programmes were considered, and a great deal of information was gathered, mostly by studying the websites. In practice, one programme was a main favourite from the very beginning: i-Tor. This is an open-source software product, developed by one of the KNAW institutes in the Netherlands (c.f. section 2.4.1). It was hardly a surprise that i-Tor was chosen as the application for this project.

During this period and the beginning of the implementation phase, the project group was supported by an external consultant from the Reekx Bureau¹², specialised in advice and support on information services. This also proved to be very useful when we had to deal with a change of project leader. Fortunately, the new project leader brought in a lot of expertise in digital data handling and archiving.

4.1.2 Implementation fase

After the content plan had been prepared and an application had been chosen, the implementation stage followed. i-Tor was a kind of 'off-the-shelf' application and available in modules, but still in development. This meant that some aspects could be rolled out immediately and that others had to wait until adjustments or bug fixing had been carried out at the i-Tor centre. In practice, this period took far more time than we had anticipated. Our programmer needed more time because an increasing number of bugs was discovered and because it was necessary to maintain contact with the i-Tor centre. Also, writing the necessary reports for the i-Tor staff about the bugs required more time. It should be stated here that cooperation between the

Fryske Akademy and i-Tor staff remained highly professional throughout the process. People were always very helpful, interested in the project, very generous with their time and showing us their goodwill.

Another aspect was the use of the OAI protocol and the creation of repositories to enable i-Tor to harvest our data. The report on these technical issues, including all other technical specifications, can be found in the Technical Report.

Website content

The CP content programme soon turned out to be very ambitious: where the CP works with criteria based on specific user groups and quality levels, far too little digital material proved to be available to employ these criteria. For that reason, we were quite happy to have at least some texts from a digital text database (Middle Frisian, before 1800, and therefore no copyright problems), a series of Mercator-Education Regional Dossiers on education in minority languages in PDF format, and one single publication of the Fryske Akademy in digital form (November 2004) in WordPerfect. This was quite a poor result for a library, but for a pilot it was just enough to test the several technical approaches for data retrieval and presentation that we wanted to explore in this project.

For the same exploratory reasons, we were very happy to find the Biblioteca Virtual Galega ready to implement a similar OAI protocol so that we could test the multi-participant approach. Also, Tresoar opened a small part of its collection in digital form to connect to DBFrysk.

In the final months of the project (until the end of April 2005), various other texts and audio-files, both from the Fryske Akademy and publisher Bornmeer, could be added to the library, increasing our experience in building a digital library.

4.1.3 Rounding-off phase

Presentations and reports

At several stages in the process, intermediate reports were produced (cf. previous paragraphs and appendices 1 – 3 and 7). Twice, a report was directly written for the KNAW. Presentations on project progress were given to a KNAW audience in Amsterdam, March 2004.

On the basis of our reports and intermediary calculations, the KNAW approved of a six-month project extension. Thus, the final date became 30 April 2005 instead of the end of October 2004. This extra time was mainly used to cover an international symposium where the project was presented to an international audience, to increase the contents of the library and finally to draw up reports and documentation.

Also, the pilot project and the intended international EU project on minority languages were presented at the DigibyB Symposium, part of the III Mercator International Symposium, in Leeuwarden on 27 November 2004. A report was drawn up on the DigibyB project proceedings from its start in August 2003 until November 2004. Project aims and purpose - in particular the Frisian pilot and the cooperation with Frisian organisations - were explained. In addition future plans relating to European cooperation were illustrated by our project leader Cor van der Meer

in cooperation with Mrs. Nieves Brisaboa, project leader of the Biblioteca Virtual Galega. The DBFrysk and EMILL websites were launched.

More presentations will follow. At the DENK / KNAW meeting on 20 May 2005, a presentation will be held called "Web Admittance: Digitalisation Projects, Web Admittances and Usability". At the end of May, a presentation will be held during the Annual Conference of IASSIST/IFDO 2005 entitled: "Evidence and Enlightenment", (May 24-27, 2005, Edinburgh, Scotland).

At several stages in the project, the advisory board or steering group ("Stuurgroep") was informed about project progress.

4.2 Project planning evaluation

Goals and actual results

The project was meant as a pilot project from the very beginning. This means that it was intended to explore the possibilities for a digital library for Frisian and other minority languages in Europe. In contrast with the DBNL, where all data are stored locally and converted into one format (both for storage and layout), this project had other ambitions, namely to investigate the possibilities for presentation from various sources in various formats from various systems and countries, and finally to offer a wider range of services such as information exchange and collaboration through the Internet.

As far as the technical goals are concerned, it can be said that the pilot has been quite successful, especially with respect to the technical integration of various types of material. Further e-services have not been implemented. Ideas about a quality-controlled and carefully planned content fill were too ambitious for this pilot. The digital library of Frisian (DBFrysk) available on the Internet is in practice not yet useful for end users, mainly because of its highly limited content.

Establishing technical links with a foreign library was originally defined as a future project extension. The fact that this has already been demonstrated at this point in time can be considered a major success.

In sum, we feel that the pilot project goals have been met. Some adjustments were made: in a negative as well as a positive sense, but nevertheless maintaining a fine balance throughout.

Evaluation of project execution and practical work

The work plan presented in the Starting Notes was indeed followed during the course of the project (cf section 4.1), including the stages discussed there.

The outcome of the various tasks fluctuates somewhat. The Starting Notes match the final outcome very closely. This means that a good job has been done there. As mentioned in section 4.1, the input from the 'Information Studio (Kennisatelier) and the advisory board or steering group ("Stuurgroep") for the practical work was limited. It turned out that more information could be gathered when this was actively sought for and when the advisory board and selected specialists were asked specific questions. Partners for practical work were also found through

personal contacts initiated by individual project group members rather than through more general approaches of spreading information about the project.

The input provided by Reekx was useful in the initial stages, considering the fact that the first project leader had little ICT experience. Reekx' contributions were particularly helpful in limiting the aim of the project, directing the discussions and bridging a period without a project leader.

The CP proved to be very useful in obtaining an idea about the difference in workload between a serious, interactive, international digital library and a pilot project concentrating on technical aspects. If the state of the contents available and the Fryske Akademy had been taken into consideration earlier in the process and if a more realistic estimation had been made of the differences between a mainly technical pilot and a fully-equipped digital library, the writing of the CP would probably not have received so much emphasis. On the other hand, a carefully written CP will certainly prove to be a sound investment once the websites and editorial boards are fully operational and functioning.

4.3 Experiences and learning effects for the Fryske Akademy

For the Fryske Akademy, this project has shown us three important things (apart from the effects connected to the core of the project):

Firstly, this project has shown us the absence of policy regarding digital material in the institute. It is quite understandable that this should be so, because until recently, a main focus always lay on the printed document or publication. These days, the printed publication is to be considered a side-product, a time stamp of a 'dynamic' digital data collection. At this moment, the Fryske Akademy is not ready for digital publishing in a web-based environment. The actual problems were discussed in sections 2.2.3. and 3.4. The need for proper policy has become very clear and will be converted into new working methods in the near future.

Secondly - as a follow-up of the first, negative experience - we definitely gained invaluable insight into the matter of organising digital processes within the Fryske Akademy: a positive development. The practical technical procedures we learned about in this process are described in the various appendices.

A third point - as a consequence of the previous points - is the awareness of the positive challenges that digital libraries and digital publishing pose to organisations. Awareness in this field has been raised considerably throughout the entire organisation. As an additional benefit, actual ICTknowledge has increased for most of the participants in the project.

The model for international use

We are grateful for the opportunity of testing the international component of this digital library concept in the pilot. The subsequent practical instructions are laid down in the EMILL document and the various technical appendices. As such, we may conclude that the concept works from a technical perspective and that it can be extended to other partners.

One aspect that has not become completely clear is the actual need for this model. The added

value it could have in gathering all this information - in many cases highly language-specific, in one extended portal or central library - is yet to be shown. By default, one may assume, with today's growing digitisation of science, that information retrieval in a web format is becoming a basic prerequisite for any scientific institute that wishes to continue functioning well. Thus, it is one of the tasks for an extended project to find partners and to define common goals on content gathering, handling and presentation aimed at a common user group. The growing pains associated with the technical realisation have already been tackled in this pilot

4.4 Project continuation / follow up

The Digibyb project was meant as a pilot aimed at a European Minority Language Library. During the project, contacts have been established in order to test the technology behind the organisational model and to work on future projects in order to make this goal feasible. The goal of this intended project is the establishment of a network of digital libraries on minority languages and the creation of a European central library on minority languages, entitled "EMILL" (European Minority Languages Library).

Despite the differences that exist between the libraries of the various language communities, this project is aimed at all of them. It will provide existing libraries with a new way of making their collections available to users, and it will also be of help to emerging and new libraries by providing them with the tools and technologies to build their own digital library according to the latest standards and technologies. Every language community will have easy access to the network through the use of common standards, an open structure and the provision of specific tools. These tools and standards are necessary to ensure participation and easy network access. The project will serve as a growing model. The objective is the creation of a large network with a central library on European Minority Languages in which ultimately all minority and / or regional languages from within the EU can be represented.

In addition to the follow-up project, this pilot also had and will continue to have a great influence on awareness levels on the part of the Fryske Akademy staff with respect to the importance not only of a digital library, but also of digital material in general. The fast growing importance and influence of the Internet is obvious. In some cases, procedures, processes and policies will have to be redefined. It is expected that this project and project team will play an important role in these changes.

4.5. Budget

The total budget for the project amounted to approximately 98,000 euro. From that amount, slightly more than 10,000 euro was reserved for external advice. Most of the remaining budget was used for hiring personnel. Less money than anticipated was required for hardware and software or for the creation of specific applications, mostly because we were able to strike a favourable deal with our colleagues of NIWI, the Netherlands Institute for Scientific Information (Nederlands Instituut voor Wetenschappelijke Informatie). They are the developers of the Open Source software tool "i-Tor". Also, spending less money in that particular period of the project, we were able to secure approval for an extension of the project deadline. The extension was duly granted by the KNAW, mainly because it could be done without any extra costs.

5 Technology

Technology has played vital role in the Digibyb project. It has led to a extensive report which will be printed separately. For the technical details, the reader is kindly referred to this separate document, entitled “Digibyb, Digital Library Project on European Minority Languages: Technical Report”.

Every local situation has its own characteristics and peculiarities, so for that reason a special technical report has been written for the Fryske Akademy in Dutch. It deals with questions how to convert our BRS databases to Dublin Core, for example, and how our OAI repositories are created and placed on the server, which server to choose, how the meta-data are harvested, and so on. Dissemination of this information is not of much use to other user groups or partners.

6 Summary

Digibyb has been a wonderful and highly interesting project. It has given us considerable and significant experiences to be used in the various fields that the Fryske Akademy and Mercator–Education operate in. The project has served as a (virtual) meeting point of experts on different levels and from different disciplines, not only for the Fryske Akademy, but also on a wider, national - and even international - level.

The Digibyb project was intended as a pilot for a European digital library on minority languages: a project designed to create a network of digital libraries and a central library on minority languages which could also serve as a portal. In addition, the project intends to assist language communities in building their own digital library. It will be clear that this future project will form a major challenge in terms of techniques employed and in terms of organisation. Most of the pilot project was carried out with this European project in mind.

In an environment with a wide array of communities and languages, it is absolutely necessary to work with open and common standards. The emphasis in the pilot has therefore been put on rather technical issues, such as searching and testing the functionality of these open and common standards.

Another project spin-off was formed by what Fryske Akademy staff learned from the various processes. Thanks to the project, it has become clear that the situation concerning (digital) publications has to change. The workflow needs to be reconsidered and in some case tasks and functions need to change. As part of the project’s outcome, a proposal will be written for the necessary changes. This will then be discussed with the management of the Fryske Akademy.

In terms of cooperation, the project has also been successful. During the project, various institutes have been willing to cooperate in order to convert part of their databases and to test the functionality of the proposed standards and techniques. Locally, in Fryslân, we worked together with “Tresoar”, the provincial library, archive and Frisian literature centre. Internationally, we worked together with the “BVG”, the Galician Virtual Library. During the III Mercator International Symposium in Ljouwert/Leeuwarden, a highly successful demonstration was given by the BVG in close cooperation with Digibyb. The BVG also intends to be a project partner in the future project “EMILL” (European Minority Language Library).

Especially in the later stages of the project, we looked more intensively for partnerships in the EMILL project. Articles have been written in newsletters and information has been sent to various potential partners. At this moment, we are engaged in talks with specialists in the Galician, Sami, Kashubian, Basque, Faroer and Catalan communities. Most of these contacts look very promising indeed.

In all, it can be concluded that the Digibyb project has been a success and that it has served its purpose very well.

7 Bibliographical references

1996

- Harter, Stephen P. - What is a Digital Library? Definitions, Content and Issues / Bloomington : Indiana University; School of Library and Information Science, 1996.

1998

- Nederlandse Basisclassificatie /red. Beheerscommissie GOO. - 3^e herz. dr. – Den Haag : Koninklijke Bibliotheek, 1998 – 148 p.
- Richvalsky, James and David Watkins - Design and Implementation of a Library – In: *ACM Crossroads Student Magazin* (1998). – 8 p.
- Van Profileren naar Samenwerking : Handreiking voor het Maken van Collectie- en Collectievormingsprofielen / [red.] Trix Bakker. – herz. dr. – Den Haag : Koninklijke Bibliotheek, 1998.

2000

- Chapman, Stephen. - Considerations for Project Management. – Andover : Northeast Document Conservation Center, 2000. – 12 p.
- Guide to Best Practice: Dublin Core, version 1.1 / Consortium for the Computer Interchange of Museum Information (CIMI) . – [S.l.] : CIMI, 2000. – 106 p.
- Noerr, Peter. - The Digital Library Toolkit - 2nd ed. - Palo Alto : Sun Microsystems, 2000. - 185 p.
- Veld, H.J. in 't - Opzet (contract) Projectdossier. – Amsterdam : KNAW, 2000.

2001

- Peereboom, Marianne and Marco de Niet - De Virtuele Bibliotheek : Definitie van Kernbegrippen – herz. ed. – Den Haag: Koninklijke Bibliotheek, 2001.

2002

- Content Management : Creëren, Beheren en Publiceren / Wisdom Assen. – Assen : Wisdom Assen, 2002.
- Eindrapport Virtueel Kenniscentrum en Expertisenetwerken / Jan Beijering ... [et al]. – Utrecht : Digitale Universiteit, 2002. – 82 p.
- Gent, Joop van and Onno Makor - Automatische Verrijking in de Praktijk – In: *Informatie Professional* – (2002) 6/7; p. 28-31
- Krol, Jelle – Frisiaca – In: PBF : *de Provinsjale Biblioteek fan Fryslân* , 150 jaar Geschiedenis in Collecties / *Jacob van Sluis, Erik Hesmerg en Petra Hesmerg* – Leeuwarden : *Tresoar*, 2002. – isbn 9080519464 – 191 p.
- Jensma, Goffe and Hans Mol. - De Nije Ensyklopedie fan Fryslân ; In Projektútstel. – Ljouwert : Fryske Akademy, 2002. – 5 p.
- Pleijsant, Jan Mark, Harriët Hangelbroek and Jan van der Starre - Information Retrieval in Luilekkerland: Slimme Zoeksystemen. – In: *Informatie Professional* . – (2002) 10; p. 30-33

2003

- Boon, Jo and Brigitte De Craene. - Procesevaluatie bij Innovatieprojecten. – In: *Onderwijsinnovatie ; juni* (2003)
- Contentplan Agriwijzer / Data Direction. – [S.l.] : Data Directions, 2003. – 57 p.
- Hillmann, Diane - Dublin Core Guide. – [S.l.] : Dublin Core Meta-data Initiative, 2003.

- Jensma, Goffe. - Notysje Digitale Publikaasjes Fakgroep SLN. – Ljouwert : Fryske Akademy, 2003.
- Aspecten aan een Digitale Bibliotheek / Ivo Zandhuis . – [S.l.] : [auteur], 2003. – 37 p.
- Zandhuis, Ivo and Ton Dijksterhuis. - Dublin Core bij het Deventer IJsselfront-project – In: *Informatie Professional*. – Vol. , no. (2003), p. 29-32.

2004

- DARE use of Dublin Core, version 1.0. / Marlon Domingus ...[et al]. – Utrecht : Stichting SURF, 2004. –23 p.

Digital bibliographical references

1997

- Day, Michael. - Notes on Mapping of National Libraries' Meta-data Requirements to Dublin Core (BIBLINK). – Bath : University of Bath ; UKOLN, 1997.
<http://www.ukoln.ac.uk/meta-data/registries/dc/biblink.html>

1998

- Savenije, Bas and John Gilbert. - Van de Traditionele naar de Virtuele Bibliotheek. – In: *Over grenzen; traditie en vernieuwing in de wetenschappelijke bibliotheek*. – Amsterdam : Otto Cramwinckel, 1998.
www.library.uu.nl/staff/savenije/publicaties/traditionele.htm

1999

- DONOR Dublin Core Gebruikersgids, versie 2.1 / DONOR Archive. – Den Haag : Koninklijke Bibliotheek, 1999.
<http://www.kb.nl/coop/donor/project-nl-index.html?coop/donor/rapporten/DCsimpleformat.html>
- Dublin Core Simple Format / DONOR Archive. – Den Haag: Koninklijke Bibliotheek, 1999.
<http://www.kb.nl/coop/donor/project-nl-index.html?coop/donor/rapporten/DCsimpleformat.html>

2000

- Information Gateways Handbook / Desire. (2000)
<http://www.desire.org/handbook/print4.html>

2003

- METS : An Overview & Tutorial (Meta-data Encoding & Transmission Standards. - Library of Congress, 2003.
www.loc.gov/standards/mets/METSOverview.html
- Powell, Andy - Guidelines for Implementing Dublin Core in XML. – [S.l.] : Dublin Core Metadata Initiative, 2003.
<http://dublincore.org/documents/2003/04/02/dc-xml-guidelines/>
- Woodley, Mary S. - DCMI Glossary. - [S.l.] : Dublin Core Meta-data Initiative, 2003
<http://dublincore.org/documents/usageguide/glossary.shtml>

2004

- DCMI Meta-data Terms / DCMI Usage Board – [S.l.] : Dublin Core Metadata Initiative, 2004.
<http://dublincore.org/documents/dcmi-terms/>
- Harper, Corey A. - DCMI Bibliography – [S.l.] : Dublin Core Meta-data Initiative, 2004.
<http://dublincore.org/documents/usageguide/glossary.shtml>

- Dublin Core in samenwerkingsprojecten en publieksgerichte ontsluiting
http://www.den.nl/standaard/dublin_core.html
- DutchESS Collectiemanagement / Koninklijke Bibliotheek
<http://www.kb.nl/dutchess/manual/colman.html>
- Introduction to XML
http://www.w3schools.com/xml/xml_what.asp
- Koninklijke Akademie van Wetenschappen (Royal Netherlands Academy of Arts and Sciences)
www.knaw.nl
- Nicholson, Dennis. - Collection Level Description
www.ukoln.ac.uk/meta-data/cld/study/collections/conspectus
- Specifications for Virtual Libraries
www.ik.fh-hannover.de/ik/person/huthloff/virtual_library/definition.htm

Internet websites of digital libraries consulted

- ACTA
http://www.acta.nl/page_nl.asp?pid=25
- Bibliopolis
<http://www.bibliopolis.nl>
- Biblioteca Virtual Galega
<http://bvg.udc.es/index.jsp>
- The Canadian Initiative on Digital Libraries (CIDL)
<http://www.nlc-bnc.ca/cidl/cidle.html>
- Die Digitale Bibliothek (DigiBib)
www.hbz-nrw.de/produkte_dienstl/DigiBib/
- EEVL (The Edinburgh Engineering Virtual Library)
www.eevl.ac.uk/
- ICODO – Project Digitale bibliotheek
<http://www.icodo.nl/nl/digibib.shtml>
- IFLA net
www.ifla.org/II/diglibtk.htm
- Libraries & Information Sciences
<http://mel.org/libraries/LIBS-michigan.html>
- National Library of Medicine
www.nlm.nih.gov/
- Virtual Library Museums Pages
<http://icom.museum/vlmp/>
- The WWW Virtual Library
<http://vlib.org/>

Publications by the Digibyb project team

2003

- Eerste tussentijdse KNAW-rapportage van het Digibybproject, Digitale Bibliotheek Europese Minderheidstalen, augustus – december 2003
- Startnotitie Project digitale bibliotheek Europese minderheidstalen / projectgroep Digitale Bibliotheek Europese Minderheidstalen; [projectleider Alie van der Schaaf]. – Leeuwarden : Fryske Akademy; Mercator-Education, 2003. – 12 p.

2004

- Applicatiekeuze voor het project 'Digitale bibliotheek Europese Minderheidstalen' : 3^e Fasedocument / Projectgroep Digibyb. – Leeuwarden : Fryske Akademy; Mercator-Education, 2004. – 6 p.
- Contentplan Digitale Bibliotheek Europese Minderheidstalen / projectgroep Digibyb. – Leeuwarden : Fryske Akademy; Mercator-Education, 2004. – 97 p.

2005

- Tweede tussentijdse KNAW-rapportage van het Digibybproject, Digitale Bibliotheek Europese Minderheidstalen, januari – december 2004 / projectgroep Digibyb. – Leeuwarden: Fryske Akademy; Mercator-Education, 2005. – 4 p.

Publications about the Digibybproject

2004

- Digitale Bibleteek Europeeske Minderheidstalen / Projectgroep Digitale Bibliotheek Europese Minderheidstalen – In: *Ut de Smidte fan de Fryske Akademy - Jrg. 38, no. 1* (2004); p. 13.
- Digitale bibleteek fan Europeeske Minderheidstalen – In: *Informatie 04 / winter 2005 .- Groningen: Reekx*, 2004. – www.reekx.nl/informatie/pdf/04.pdf
- Digibyb – In: *Reposi-Tor*. – No. 4 (2004). – www.i-tor.org/nl/reposi_tor_i_tors_elektronische_nieuwsbrief/archief_reposi_tor/reposi_tor_nr_4_juni_2004/
- Digitale Bibleteek Europeeske Minderheidstalen / Projektgroep Digitale Bibliotheek Europese Minderheidstalen. – In: *Ut de Smidte fan de Fryske Akademy – Jrg. 39, no. 1* (2005); p. 10

Presentations concerning the project by members of the project group

2003

- Presentation Digibyb project, October 2003, Leeuwarden, by Alie van der Schaaf (Dutch).

2004

- Presentation Digibyb project, KNAW-meeting organised by DENK / KNAW, March 10, 2004, Leiden, by Elly Albers (Dutch).
- Presentation Digibyb project at the IASSIST conference, Madison, USA, May 2004 by Cor van der Meer (English).
- Presentation concerning the Digibyb project at the Minisymposium of the Social Sciences Faculty of the Fryske Akademy, 16 May 2004 by Elly Albers (Dutch).

- Presentation Mercator-Education project and the Digibyb project at an international meeting in Brixen / Bressanone, Italy, 4-7 November 2004 by Cor van der Meer.
- Presentation Digibyb project; pilot Frisian at the Digibyb symposium, 27 November 2004 ; III Mercator International Symposium, 25-27 November 2004 / Cor van der Meer
- Presentation BVG, Digibyb- and EMILLproject; European Minority languages Library website at Digibyb symposium, 27 November 2004; III Mercator International Symposium, 25 – 27 November 2004 / by Cor van der Meer and Nieves Brisaboa (Galicia, Spain)

2005

- Presentation Mercator-Education project and the Digibyb project during the official visit of representatives from the Hungarian Parliament on 3 February 2005 to the City of Leeuwarden/ by Cor van der Meer, Tseard de Graaf and Tsjerk Bottema.
- Presentation Digibyb project at the DENK / KNAW meeting called: Web admittance; digitalisation projects, web admittances and usability/ by Cor van der Meer, May 20, 2005, Den Haag.
- Presentation at the IASSIST/IFDO 2005 Annual Conference: Evidence and Enlightenment, 24 – 27 May 2005, Edinburgh, Scotland, by Cor van der Meer

Visits

2003

- Visit Digitale Bibliotheek voor de Nederlandse Letteren (www.dbnl.org), Leiden, 18 March 2003 / project leader Cees Klapwijk.
- Visit NWO, Den Haag, 23 April 2003, project SHARED / project leader Agnes Suarez.
- Visit of Ivo Zandhuis to Mercator-Education / Fryske Akademy, 16 April 2003.

2004

- Visit of Laurents Sesink (i-Tor) and Joris van Zundert (BNTL+) to the Fryske Akademy / Digibyb project, March 2004.
- Visit to i-Tor, June-July 2004 by Digibyb project team.

Attended congresses / workshops / courses

2003

- Report on the visit to i-Tor workshop on October 30, 2003 / Bob Boelhouver (Dutch).
- Report NVB-conference, Ede, on November 3, 2003 / Elly Albers (Dutch).
- Report DENK / KNAW meeting “Bibliotheek en Informatie” on November 19, 2003 / Peter van der Meer.

2004

- Mini-seminar Taxonomie / Reekx on February 25, 2004 / Elly Albers
- On-Line Conference, April 2004, Amsterdam by Digibyb project team.
- Report NVB-SB Conference, June 10, 2004 / Elly Albers.

(Footnotes)

¹ What are digital libraries? / Donald J. Waters – In : *CLIR Issues*; no. 4 (1998); p. 1-10

² www.fryske-akademy.nl

³ www.tresoar.nl/

⁴ <http://bvg.udc.es/index.jsp>

⁵ www.dbnl.org

⁶ www.dbnl.org

⁷ www.bornmeer.nl

⁸ www.mercator.org

⁹ www.mercator-education.org

¹⁰ www.i-tor.org

¹¹ www.mercator.org

¹² www.reekx.nl