First preprint, August 2001
Revised version, February 2003

Sagnanet – An on-line Image Access to Rare Materials
Overview Across Three Collections – Access and Usability

Kristrún Gunnarsdóttir, kg62@cornell.edu
Cornell University Library, Division of Digital Library and Information Technologies
Ithaca NY, USA

• Abstract

The Sagnanet on-line database consists of the Icelandic family sagas, Germanic/Nordic mythology (the Eddas), the history of Norwegian kings and tales from the European age of chivalry. In the beginning of the year 2001 the database and the interface design was at a final stage of development, and most of the material already available on-line. A usability study was conducted between January 25th and June 17th, 2001, focusing on a number of user groups and general user behaviour.

It will be argued in this report that the geographical scope of visitors partaking in the study and the number of unique *items*[1] that where downloaded, indicates an importance of this service for at least preliminary research, however, more in-depth studies still require hands-on contact with the items and a direct access to local expertise at the collections that hold them. The field of study is by no means popular. Presumably there are less then five hundred researchers worldwide in the areas of medieval Nordic literature, Old Norse, Icelandic cultural history of the medieval and late medieval periods.[2] Nevertheless, it will be concluded that the use of this on-line repository indicates an increase in demand, which is exclusively generated by offering the material in this format.

The report to follow outlines step-by-step preparations and the technical implementation of the study. The two chapters *Summaries* and *Discussion* portray statistical results followed by discussion of key concerns and conclusions.

NOTE! Appendices are 125 pages and available at: **http://sagnanet.is/Appendix_a4.pdf**

[1] **Item**: Throughout this report *item* will be used for any single manuscript or book even if they are usually referred to as multi-item digital objects (holding multiple image items (page=image) linked to one object identifier).
[2] This figure is arrived at by an examination of the Sigurður Nordal inventory of scholars worldwide and their database of institutes across the world that teach Icelandic or Old Norse, also by general querying at the institutes that the listed scholars work for.

- **Introduction**

The Sagnanet on-line database was under construction for over four years. This is a large-scale digitisation of about 250.000 manuscript pages from the National and University Library of Iceland (hereafter NULI), the Árni Magnússon Institute in Reykjavík, Iceland (hereafter AM or AM Institute), and The Fiske Icelandic Collection at Cornell University Library, USA (hereafter FIC), a collection that contributed 150.000 printed pages (ca. 150 books). In collaboration, the three institutes produced a unique virtual collection that is currently accessible on-line.[3]

The Sagnanet on-line collection consists of the entire range of Icelandic family sagas. It also includes a very large portion of Germanic/Nordic mythology (the Eddas), the history of Norwegian kings, so-called contemporary sagas and tales from the European age of chivalry. A great number of the manuscripts also contain Icelandic ballads, poetry or epigrams.

To provide efficient access to the image content a team of cataloguers and literary experts catalogued 650 manuscripts in the NULI holdings. The catalogue is very detailed and gives an account of each item's list of contents. This accuracy and detail in the catalogue was seen as necessary given the nature of the NULI collection. Manuscripts in the NULI holdings are poorly researched and most of them are some sorts of compilations, sometimes of wholly unrelated materials. A decision was made to always scan each item or volume from beginning to end, and the detailed catalogue is a tool for users to find what is inside.

The AM catalogue was not in digital format when the Sagnanet project started, but the Institute holds a card catalogue, that was updated and digitised for the purpose of Sagnanet. AM manuscripts within the material scope of Sagnanet are very seldom compilations, and if so, it is always of closely related materials.

The Cornell University catalogue holds data about the approximately 150 books that are a part of Sagnanet. An extract of the records from the FIC catalogue was combined with the Icelandic catalogues and they made searchable across all three partaking collections with a Z39.50 search protocol. During the design stage of the interface, a group also designed direct access to the image database from various perspectives, experimenting with how users could get a sense of overview by browsing the database content. The result was a number of functions to browse through titles proper or uniform titles, a subject headings list and a browse through so-called 'collective supplied titles' for many of the compilations in the NULI holdings. The extensive cataloguing of uniform titles provided the key for an overview across all three collections. Some of the metadata accompanying the FIC books needed to be altered for this purpose and presently close to 800 uniform titles are accessible on the site.[4]

[3] http://sagnanet.is

[4] See also the Sagnanet website, *Information about the project – Scope – Technology*

The following is a step-by-step description of the usability study. It was one of the outlined goals of the project, that such a study would be conducted*, focusing on the advantages and disadvantages of working with this material by having sole access to it through the Internet*.[5] We invited a number of groups to be *registered visitors* on the condition that their use of the site would be monitored. After four months of monitoring our visitors were asked to evaluate some basic key features of the site.

- **Organizing the study**

User groups

From the onset, three groups were self evidently prospective user groups and lists of them are readily available: The **Sigurður Nordal Institute's Inventory of scholars,** a database containing information about scholars in the field of Icelandic Studies world wide, and **Icelandic World Wide,** a database containing information about the teaching of Icelandic in Universities across the world including links to the institutes that teach Icelandic.[6] These inventories cover groups of scholars and researchers in the following categories:

1. Medieval Nordic literature
2. Old Norse and comparative linguistics
3. Icelandic / Nordic cultural history

The documents were used as the main guide for targeting users. They list individual scholars and researchers in association with the institutes and universities they work for. The area of their particular research interest is listed, and how they can be contacted.

The two inventories provided a good basis, both for choosing from them and for searching for scholars and researchers in other but related areas at the institutes and universities on those lists. Additional areas of study are:

1. Fine art, graphic design, typography
2. Paleography
3. Information studies - Library science

[5] The Sagnanet website, *Information about the project -Goals*
[6] http://www.nordals.hi.is/engindex.html

Primary rule for choosing was to pick from the inventories all that use email. This choice was naturally arrived at because of our plan to use the electronic network to contact people, but more importantly because Sagnanet is an Internet repository and would in any case require users that are already networked and have acquainted themselves with Internet resources, at least email and the most basic information resources. Scholars (in the other but related areas of study) were chosen randomly from various places in the world. An invitation letter was sent out. The scope and goals of the project were introduced and scholars and researchers encouraged to directing their graduate students to the site for research purposes. [**Appendix A**]

Over the time period from January 25th until June 17th, 107 registered visitors accessed the site from 179 machines, hitting the site from altogether 18 identified countries. The number 107 is limited to visitors that accessed more then 10 sources, which is set here as the minimum to be considered as having seen the structure of the site and recognise what it offers. 11 visitors accessed 150 or more sources, 26 visitors 50 or more sources, and 40 visitors accessed 30 or more sources. These figures will be outlined further in the *Summaries and Discussion* chapters but just to give an idea of the scope of visits, the total number of machines accessing the site was 295.

User registration

The design of the '*First Time Visitor Registration*' page was fairly simple with no required fields beside visitor chosen username and password. We only needed a very general idea of *who* the visitors are. Most all left their email address, which connected them already with the user list that this study started out with, i.e. if they were on that list to begin with. The goal of user registration was primarily to make possible the monitoring of individual auto-generated user IDs and track individual *user histories*. Identifying the persons was irrelevant.

The registration form shows that a variety of visitors were expected to use the site, not merely scholars and professionals and it was interesting to see if people that simply had layman interest in the material or students of lower educational levels, would come across this material and make use of it. It was also important that visitors were aware of the monitoring.

To monitor the site and the use of the database gave an opportunity to see exactly how visitors traverse pages, what they are looking at, and in what order. It was for example important to see how much our visitors use the *user free-text input* search mechanism versus *browsing options* that only give 'canned' overviews of what is available. Another important issue was how much visitors want to know about the project, its scope and goals, or how much they like to see packaged introductions to the material, compared to how much they

actually access the image material. Still another important issue was the usefulness of help pages and how much they are accessed relative to use of image views. [**Appendix B**]

Automated data collection

The technology used for monitoring visitors and collecting data about their visit is based on a custom designed web server log. What goes into the log is limited to the accessing of web pages that talk directly to the Sagnanet database. A single log would contain the following:

1. IP number of machines visiting
2. User IDs accessing
3. Date time of request
4. What web page is visited
5. What items are being viewed (manuscripts and books)
6. The language of the interface (English or Icelandic)

At the top of the list is the *IP number* of all incoming machines for identification of the countries they come in from. In theory this should be sufficient to see the total number of visitors and identify their origin, but this sort data is not accurate.

a) A single visitor can be situated in an office or an institute that supplies his or her machine with a new IP number every time the desktop computer is restarted. An option in that case is to identify the IP range but then again there might be more then one visitor accessing from that same range.
b) Not all machines are registered domain names, even if they have IP numbers and operate normally on networks.

This is shown in the summary chapter. However, the collection of IP numbers is an important tool to get a picture of the geographical scope of visits over the whole time period and even more importantly to see changes from one month to another.

Second and third on the list are *user IDs* and *date and time* to see each visitor's user history on a timeline.

Fourth is the logging of every request for a *web page*. The web page names identify what requests are being made into the database. This makes it possible to recognize and distinguish between requests for help functions, for pages that give overview of database content, for use of free-text search, tables of content for individual items and image pages.

Fifth and sixth are identifications of *individual items* (manuscript or book IDs and/or titles) and whether the *language* of the interface is English or Icelandic. [**Appendix D**]

Individual evaluation

Sagnanet guests, who registered their email addresses where contacted again before the end of May and asked to evaluate the features of Sagnanet thought to be crucial to outcome of the finished product: [**Appendix C**]

1. Navigability: finding one's way through the Sagnanet site.
2. Clarity and effectiveness of search and browse options: finding specific items or arrays of items.
3. Usefulness of help functions and clarity of help information.
4. Usability of images for reading and research.
5. Usability of images for teaching.
6. Usefulness of accompanying data (cataloguing information and view of contents).
7. Sagnanet as an information resource for Old Icelandic and Nordic medieval studies.
8. Sagnanet as an information resource for historic and cultural research.
9. Sagnanet as an information resource for artistic studies (typography and/or illustrations in books and manuscripts).
10. Sagnanet as an information resource for other educational purposes and for public awareness.

Interestingly, we did not get much response to the questionnaire. Almost a hundred of Sagnanet guests visited the page, but only 21 took the time to evaluate the features of Sagnanet listed for them. Nevertheless, the questionnaire page would feed user input directly into a database, into which the web server log was also dumped. From there an access to the results was provided for analysis, locking together individual user IDs and user histories, and evaluations if applicable. Many visitors were also eager to express them selves in writing, which provided a valuable additional resource.

Data processing

Database design and programming was needed to manage the usability study so as to minimize the labour of organizing and analysing the data. The following summarizes the set up:

1. An Apache web server collecting all user-history-data into one log file. [**Appendix D**]
2. *Perl* programming used to analyse the log file.

3. Usernames created by visitors converted into unique identification numbers.
4. A relational database environment *INFORMIX* designed to host the data output of the program. [**Appendix E**]
5. An automated input from the questionnaire linked to the data structure.
6. Standard queries accompanied with simple arithmetic functions made to extract views onto the data structure as needed.

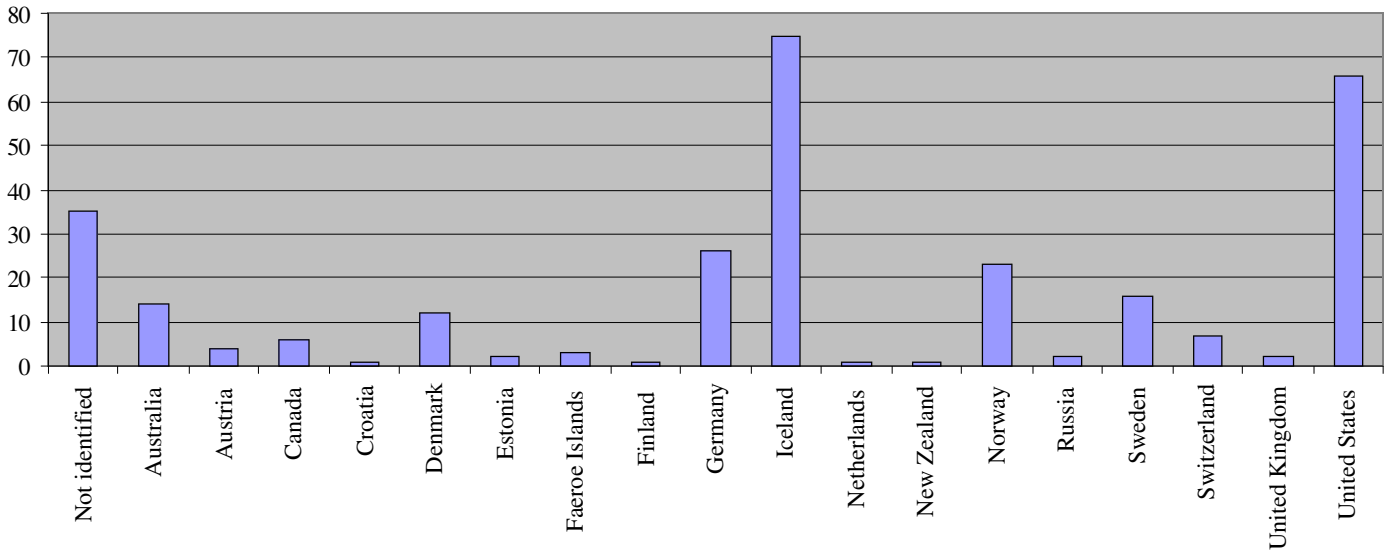
- **Summaries**

This chapter will portray multiple views onto the data that was generated over the time period and divided into seven categories:

1. Geographical scope of visits over the whole time period
2. Variation in frequency and scope from month to month
3. The geographical scope viewed with reference to visitor status
4. The geographical scope and visitor status with reference to usage
5. Using the site – Accessing the image content
6. Individual user histories – individual evaluations
7. Interface structure – design issues

Geographical scope of visits over the whole time period

The total number of incoming machines was **295**, from **18** countries in **3** continents. Iceland and the United States had the highest number, *75 from Iceland* and *65 from the United States*. Countries in a medium range were *Germany* 26, *Norway* 23, *Sweden* 16, *Australia* 14 and *Denmark* 12. Other countries had an average of about 3 incoming machines each. It is noteworthy that 35 machines are not geographically identified.



(Fig. 1)

Variation in frequency and scope from month to month

The geographical scope of visits and the number of identifiable machines varied somewhat from month to month. The table below shows how this happens. Total number of identified machines is highest in the first month and drops significantly during the next two. However, from April 25th until the end of the study, the number increases largely again. The geographical scope however, is greatest in the second last period, when visits are from 14 countries, but they are from 12 countries in the first month.

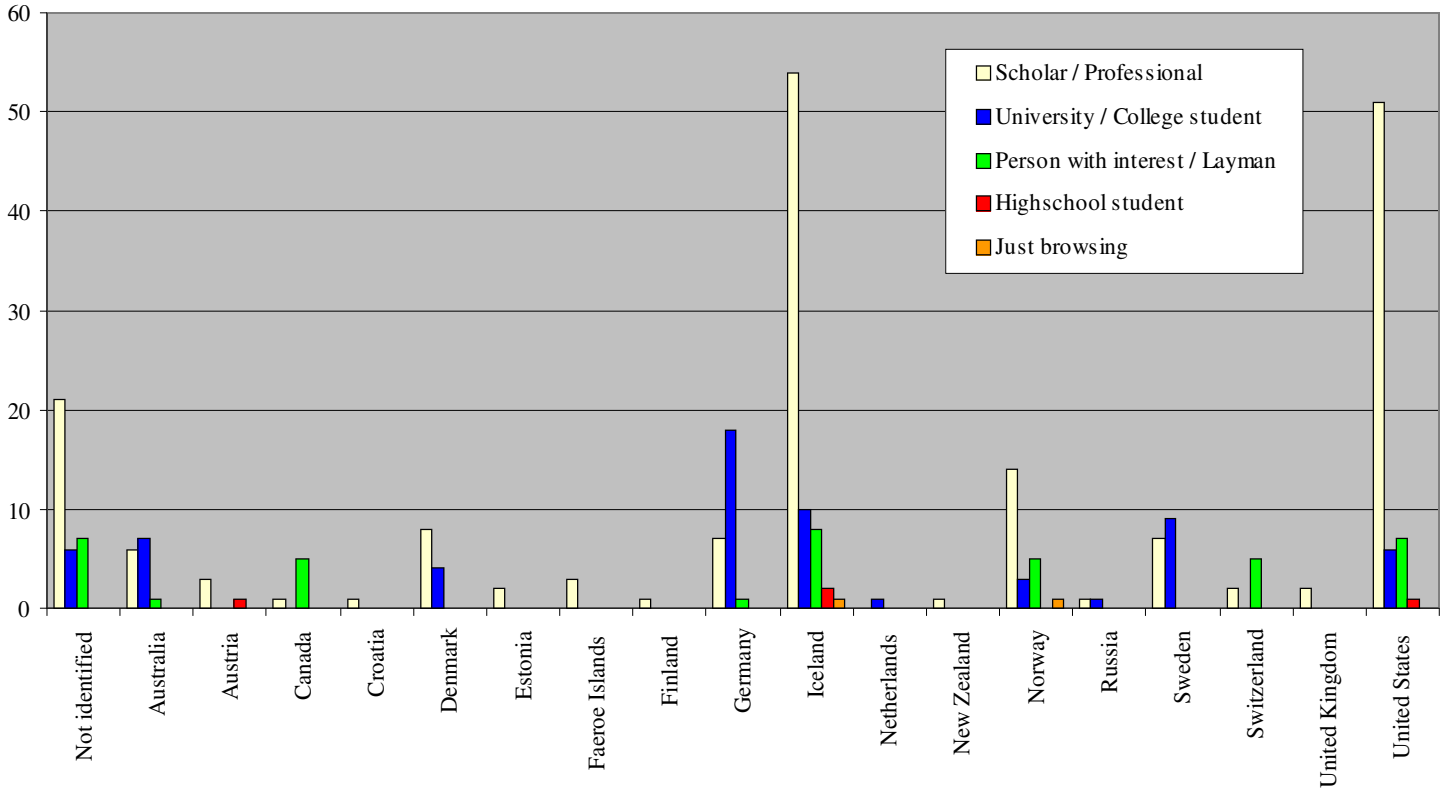
The table in Fig. 2 does not show *how often* each machine visits only *how many* unique are identified over a month at a time. Fig. 3 shows uniquely identified machines over the entire period. Comparison of the two shows that access is distributed over more then one month's period and indicates that a number of visitors returned to the site. *User histories* here below give more detailed account of individual usage.

Jan. 25 – Feb. 24		Feb. 25 – Mar. 24		Mar. 25 – Apr. 24		Apr. 25 – May 24		May 25 – Jun 17	
Not identified	8	Not identified	2	Not identified	6	Not identified	10	Not identified	12
Australia	4	Australia	3	Australia	1	Australia	6	Australia	2
Austria	2	Austria	1	Austria	2	Austria	1		
						Croatia	6		
		Canada	1			Canada	1		
Denmark	3			Denmark	1	Denmark	1	Denmark	9
				Estonia	2				
Faeroe Islands	1	Faeroe Islands	1			Faeroe Islands	1		
						Finland	1		
Germany	11	Germany	5	Germany	2	Germany	8	Germany	3
Iceland	32	Iceland	6	Iceland	17	Iceland	23	Iceland	16
								New Zealand	1
Netherlands	1								
Norway	2	Norway	2	Norway	7	Norway	10	Norway	5
Russia	2								
Sweden	5	Sweden	4	Sweden	5	Sweden	3	Sweden	4
						Switzerland	1	Switzerland	7
United Kingdom	1					United Kingdom	1		
United States	34	United States	3	United States	2	United States	17	United States	12
12	106	9	28	9	45	14	90	9	71

(Fig. 2, bottom row shows total number of countries and total number of machines over each period.)

Country	Number of machines
Not identified	34
Australia	14
Austria	4
Canada	6
Croatia	1
Denmark	12
Estonia	2
Faeroe Islands	3
Finland	1
Germany	26
Iceland	75
New Zealand	1
Netherlands	1
Norway	23
Russia	2
Sweden	16
Switzerland	7
United Kingdom	2
United States	65
Total	**295**

(Fig. 3)



(Fig.4)

The geographical scope viewed with reference to visitor status

Six countries presented only visitors of the *Professional / Scholar* status. These are Croatia, Estonia, Faeroe Islands, Finland, New Zealand and the United Kingdom. Only a few registered visitors were coming in from these countries, a total of 10, on the average 1,67 --or less then 2 per country. [**Appendix F**]

Professional / Scholar status visitors are overall the majority of visitors, a total number of 185. 11 countries present over 50% of their visitors with this status, Austria (3), Croatia (1), Denmark (8), Estonia (2), Faeroe Island (3), Finland (1), Iceland (54), New Zealand (1), Norway (14), United Kingdom (1) and USA (51).

The *University / College student* status is the second largest group, a total of 65 visitors, and the majority of visitors from 4 countries hold this status, Australia (7), Germany (18), Netherlands (1) and Sweden (9).

39 visitors registered as *Person with interest / Layman*. They come in from Australia, Canada, Germany, Iceland, Norway, Switzerland and USA, thereof are Canadian and the

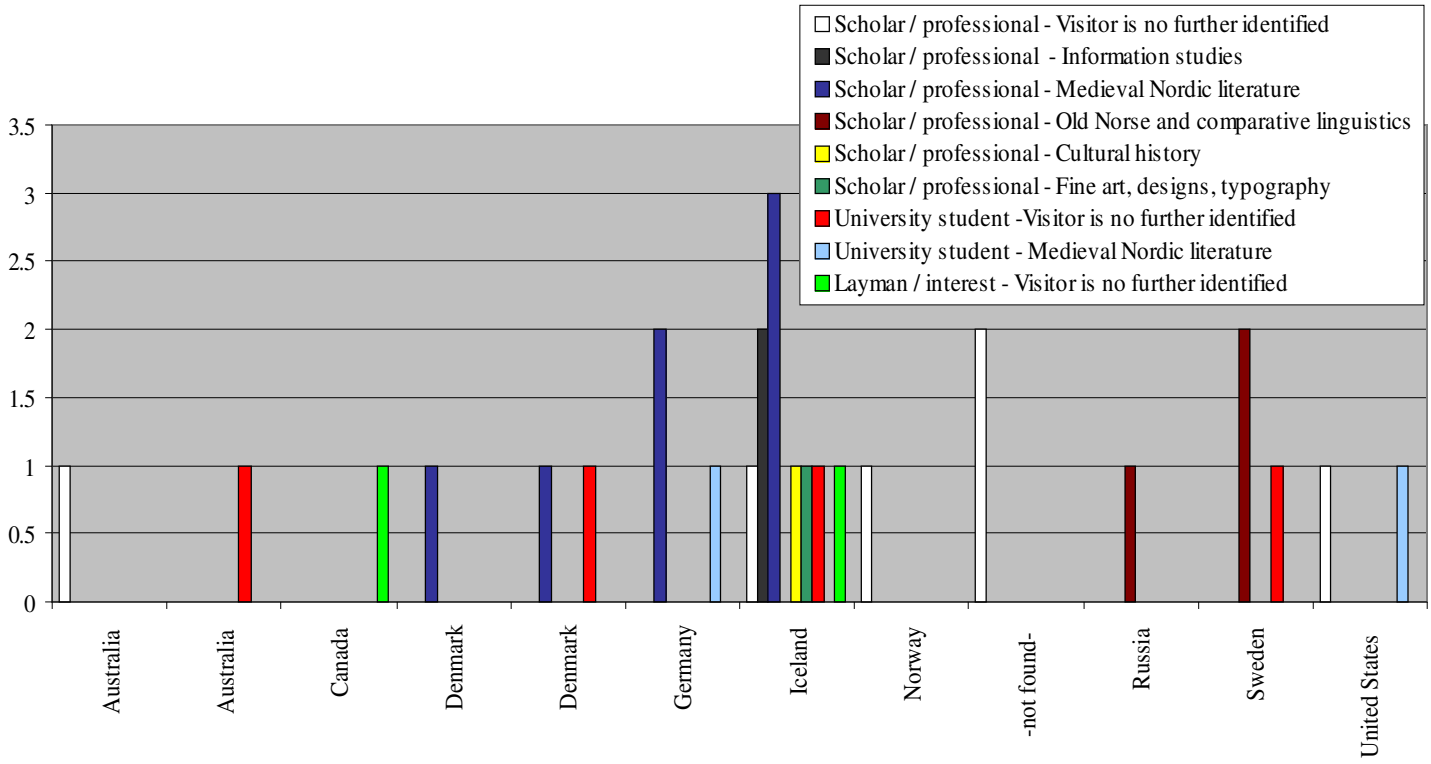
Swiss visitors with this status, 5 from each country, 83.33% of the total number of Canadian visitors, and 71,43% of the total of Swiss visitors.

4 *High-school students* registered, 1 from Austria, 2 from Iceland and 1 from USA, and 2 registered visitors claimed they were *just browsing*.

The geographical scope and visitor status with reference to usage

The chart in Fig. 5 shows registered visitors that visited 50 or more pages. Many of them made extensive use of the Sagnanet database of the time period, which is outlined in Appendix J. The total number is 28, thereof are 2 visitors not located geographically. The largest group is scholars in medieval Nordic literature (7). University students are 5, thereof 1 identified specifically as studying medieval Nordic literature. 2 with that identity appear on the chart, but were identified as the same visitor, i.e. the same user ID (21), coming in from both Germany and USA. [**Appendix G**]

Scholars and professionals with no further identification are six, from at least 4 countries, Australia, Iceland, Norway and USA. 2 of them show up with unidentified locations, but one of those two was recognized to be one of the visitors from USA, i.e. the same user id (26) showed up. 3 visitors focusing on comparative linguistics are registered, 2 in Sweden and 1 in



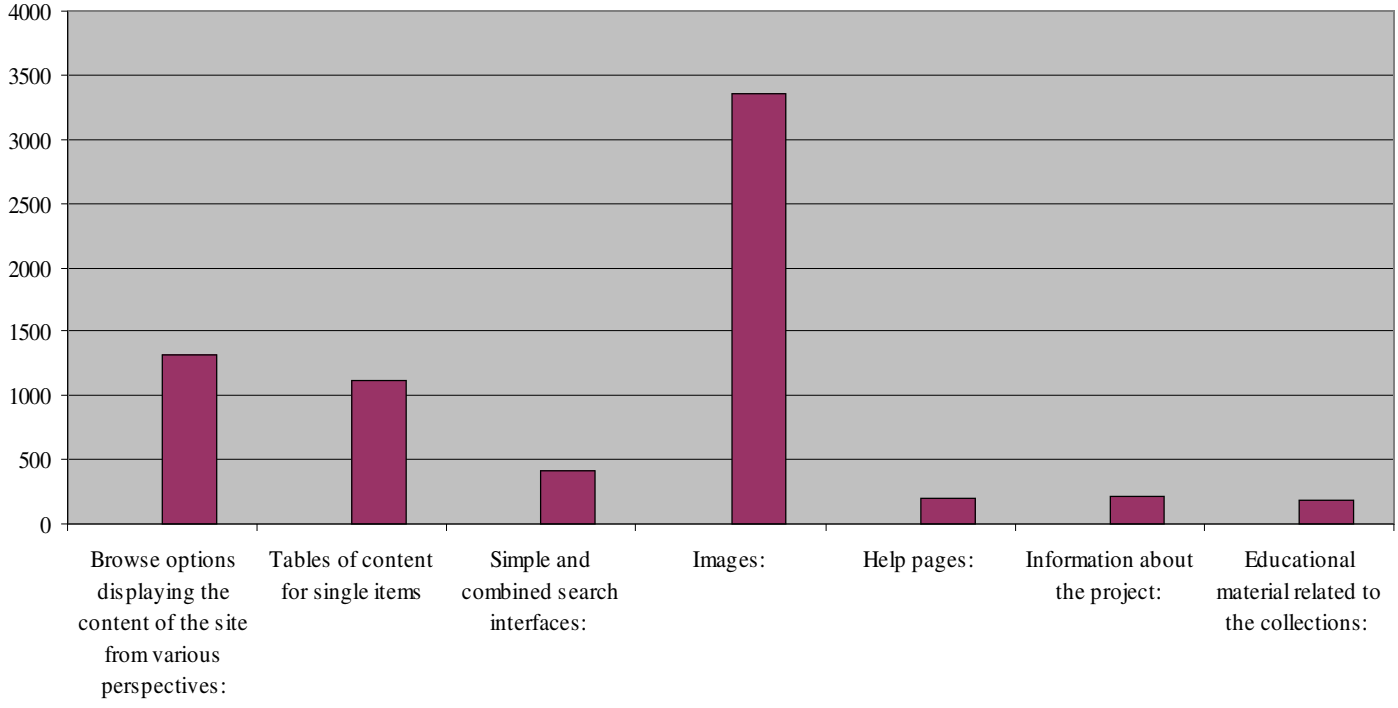
(Fig. 5)

Russia. 2 visitors focus on information studies, 1 on cultural history and one on fine art & designs, all based in Iceland. 2 laymen are registered, one in Canada and one in Iceland.

Using the site – Accessing the image content

Fig. 6 shows total visits to web pages by categories. The total number of web pages visited over the time period is 6821. All web page figures are restricted to pages that talk directly to the database, and are accessed by users that visited at least 10 such pages.

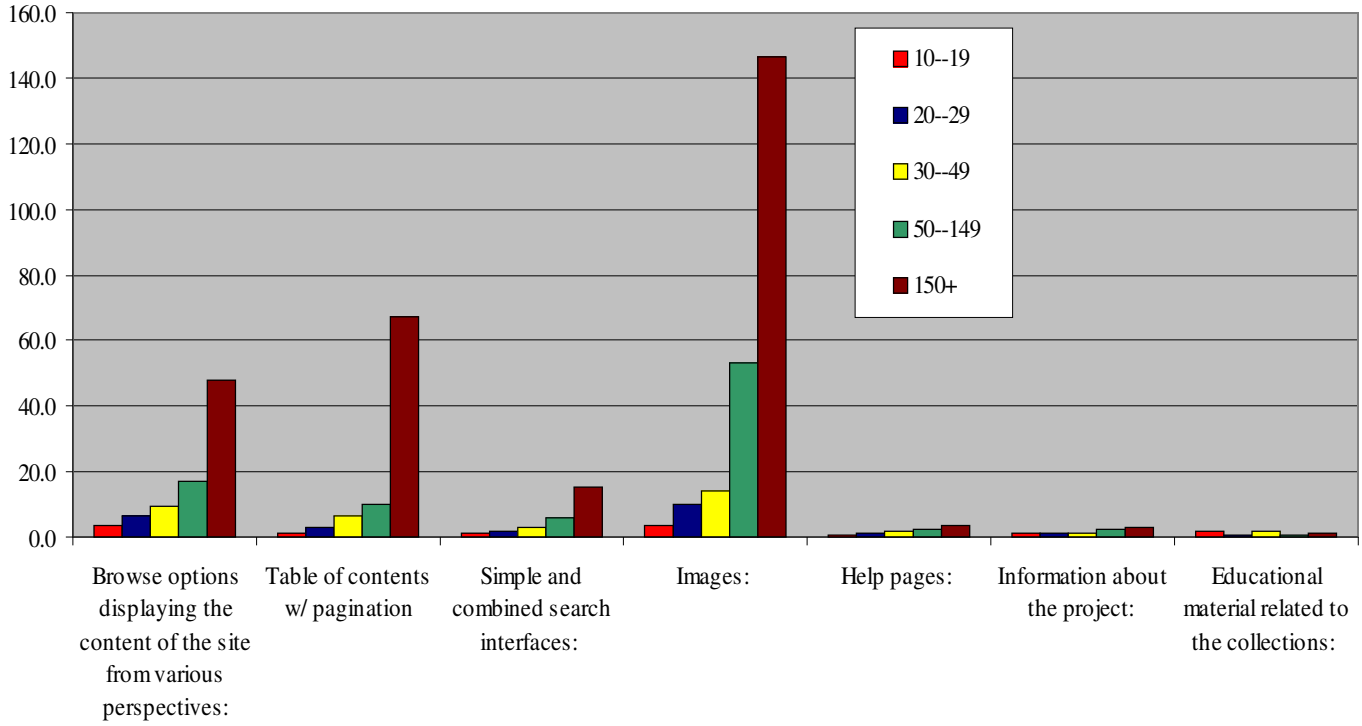
Visits to pages that provide overview of the site's content are 1325. Use of content tables for individual items is the total of 1122 visits. Use of simple and combined search interfaces is a total of 414 visits. Total number of recorded image page projections is 3355.[7] Visits to help pages are 200. Visits to information about the project are 222 and visits to educational material related to the Sagnanet collection are 183. English interface was used in 38.81% cases and the Icelandic interface in 61.19% cases.



(Fig. 6)

The chart in Fig.7 shows the relationship between how visits distribute at the site, and the average total number of pages visited.

[7] This figure may be inaccurate, because at a closer look at the webserver log thumbnails are not always recorded (should be 5 for each thumbnail view) and if visitors are impatient, a full view page may record without an image being recorded.



(Fig.7)

Pages that contain support material and give an overview of the site, are such as subject headings lists, an option to browse by uniform titles or titles proper, and to browse items that have been supplied with collective titles. The use of these browse options compared to use of the free-text search is in the ratio **48,0 / 15,6** for 150+ visited pages, **17,2 / 6,0** for 50-149 visited pages, **9,3 / 3,1** for 30-49 visited pages, **6,3 / 2,0** for 20-29 visited pages and **3,7 / 0,9** for 10-19 visited pages. Furthermore, tables of content are heavily used and quite often accompanied with the access those tables give to all other items (across the 3 partaking collections) that contain any of the *uniform titles* that may appear along with titles proper in a content table. [**Appendix H**] A growth in image use per free-text search is detected if the total use of images is measured against free-text searching.

visited pages	**avg. use of *free-text* search**	**avg. image use per ID**	**img/per *free-text* search**
10-19	0,9	3,8	4,22
20-29	2,0	10,2	5,10
30-49	3,1	14,3	4,61
50-149	6,0	53,0	8,83
150+	15,6	146,3	9,38

(Fig.8)

A similar growth of image use is detected in relation to the use of browse functions and other overview pages. However the image number per browse option is much less here.

visited pages	avg. use of browse options & overviews	avg. image use per ID	img/per browse & overview
10-19	3,7	3,8	1,03
20-29	6,3	10,2	1,62
30-49	9,3	14,3	1,54
50-149	17,2	53,0	3,08
150+	48,0	146,3	3,05

(Fig. 9)

This becomes clearer when taking a closer look at the ratio between the average use of free-text search and the average use of browse options and other overview pages.

visited pages	avg. use of user input mechanism	Avg. use of browse and other overview functions	ratio 1:
10-19	0,9	3,7	4,11
20-29	2,0	6,3	3,15
30-49	3,1	9,3	3,00
50-149	6,0	17,2	2,87
150+	15,6	48,0	3,08

(Fig.10)

The ratio ranges from ca. 1:3 to 1:4, which is was a surprise to the Sagnanet design team. The free-text search is very effective and offers both simple and combined searches, and much work was invested in the improvement of it. Nevertheless browse and other overview options were also carefully thought out even if they are technically simpler, and seem to be very straightforward and easy to use for visitors. It might be suggested that some overviews like that, are felt like short cutting to the image material and furthermore as providing a *true* sense of what exactly the database contains. It may also be that visitors are reluctant to use free-text search because of uncertainties with spelling and special characters.

Growth in use of the general help page and help w/ search is steady and approximately the same until visits exceed 150+ web pages. Then the use of the general help page takes a jump, but not the use of help with search. The latter is a detailed description of search and browse functions but the former gives only a very general idea and refers to the latter document.

Visited pages	10-19	20-29	30-49	50-149	150+
Help with search	0,1	0,3	0,4	0,5	0,6
Help using the Sagnanet website	0,7	1,1	1,4	1,7	**2,8**

(Fig. 11)

The sudden growth in use of content tables of individual items when the average visit exceeds 150 pages per user ID is also noteworthy but this could be explained.

Visited pages	10-19	20-29	30-49	50-149	150+
Table of contents w/ pagination	1,3	2,7	6,6	9,9	**67,5**

(Fig. 12)

It is recognized from the raw data that individual user sessions never last for long and both the environment and some of the material is new to the users, and therefore it can be argued to some extent that the use of content tables and extensive use of browse and other overview options, is a sign of a newcomer's behaviour in the on-line environment. Sagnanet visitors seem to be traversing the site in order to *learn* what is there. It could be argued as well, that the most use made of this environment over a period of only 4 months and 3 weeks, is by serious researchers who should be already familiar with much of the material content. They are also likely to be informed about the scope and goals of the Sagnanet project. Their behaviour *can* therefore be interpreted in such a way that they make use of content tables in relation to this new and unique overview possibility, that the three partaking collections can be seen in a relational context. The '*Items containing same uniform title'* page can be arrived at both through drop down menus containing uniform titles and also directly through content tables.

A new insight is provided to the material physically residing in the collection of the National and University Library of Iceland. The NULI collection has never been much researched. It was mentioned in the introduction here that a total of 650 manuscripts were catalogued in great detail, solely for the purpose of this project. This has resulted in many large displays of content tables for single NULI items. Now the collection in conjunction with the printed books of the Fiske Icelandic Collection at Cornell University can be seen in

context with the older and authoritative manuscripts of the Árni Magnússon Institute in Reykjavík.

Individual user histories of 150+ visited pages support this interpretation of user behaviour. As an example, user **ID 47** takes a look at 3 NULI manuscripts and 3 AM manuscripts all containing the uniform title: *Eyrbyggja saga.* User **ID 208** looks at 3 books from FIC, 2 from AM and at least 5 from NULI, in search for genealogical material and registry residing in the uniform title *Íslendingabók*, or other material related to the Icelandic settlement. User **ID 21** makes use of overview possibilities across all three collections in search for *Icelandic family sagas* in 9 FIC books, 12 AM manuscripts and 1 NULI manuscript. User **ID 150** looks at 5 FIC books and 2 NULI manuscripts containing the uniform title *Eddukvæði*, and also looks at 6 NULI manuscripts and 4 AM manuscripts, all given a collective supplied title, *Sögubók* or Storybook, containing mixtures of family saga literature. User **ID 12** looks at 4 NULI manuscripts, 4 AM manuscripts and 3 FIC books, all containing the uniform title *Flóamanna saga*. User **ID 60** looks at 8 AM manuscripts and 2 NULI manuscripts, all containing the uniform title *Kjalnesinga saga*. [**Appendix J**]

A excessive use of content tables, especially when visits exceed 150 pages average, can also be associated with long lists of items appearing in search or browse results. The user is checking the contents of many items at once. It seems though like in one case at least in the 150+ group, that user's impatience causes a repeated reload of the same content table. All user histories in the two largest groups make use of content table w/pagination to a large extent directly in conjunction with a full image view. Often the '*Items containing same uniform title'* page and a full image view are lined up together, with or without an associated content table view, and accessing uniform titles is almost exclusively a tool to move between items containing the same material, whether or not this is done across the three collections.

According to figures in the Sagnanet economic analysis document[8], 1200 loans were registered at the Department of Manuscripts at NULI in the year 2000. It is estimated that 10-20% thereof, is manuscripts containing material within the material scope of Sagnanet, which gives a maximum figure of 240 items. In the same year, the AM Institute in Reykjavík reports 43 visitors, accessing manuscripts in their collection. Typically, about 300 books a year are borrowed from the FIC. No further records are kept of quests for books, but if it is estimated that 20% is within the Sagnanet material scope, in parallel with estimates at NULI, the total of yearly quests for Sagnanet material at FIC is 60 books. Compared to these figures that come to a total of roughly 343 accessed items in a year,[9] the total number of accessed

[8] *Open Access to a Cultural Heritage: An Economic Analysis of the Sagnanet Project*, Sveinn Agnarsson, Institute of Economic Studies, University of Iceland – June 2001, http://sagnanet.is/EconomicStudy.pdf

[9] It doesn't say exactly how many AM visitors were researching material within Sagnanet material scope in the year 2000. It can be estimated though that most all AM visitors are limited to a very narrow field of study, for example of a single saga or a philological study of a singe or a few manuscripts.

items via the Sagnanet on-line database, over a period of 4 months and 3 weeks, was 512 manuscripts and books. Thereof are FIC books a total of 91. The total number of images viewed in *full view,* is 2523. 238 *thumbnail view* pages were displayed and 59 thumbs quickly enlarged. 126 full size images where zoomed at and the estimate of 117 images printed.[10]

Individual user histories – individual evaluations

Registered visitors, accessing 50 pages or more are giving a good indication of how a site like Sagnanet *can* be used. Whether or not this is how the site *will be used* is unsettled. User histories indicate that the site is ideal to search for material and look briefly at various items, and in some cases visitors engage in a lengthy viewing of single items. Examples of that are user IDs **47**, **12**, **60**, **85**, **21**, **166**, **11**, **109** and **178**. [**Appendix J**] It is still difficult to say though what users of the two most active groups are thinking. Only a few of them took the time to reply to the questionnaire but those who did also took the time to write constructive and critical comments.

User IDs **47**, **21**, **85**, **120** and **131** from the 150+ group responded, 5 out of 11 --or 45,45% of the group and 25% of all response. User IDs **93**, **153** and **204** responded from group 50-149, 3 out of 15 --or 20% of the group and 15% of all response. User IDs **87, 105** and **176** responded from the 30-49 group, 3 out of 14 --or 21,43 % of the group and 15% of total response. IDs **1** and **42** responded from the 20-29 group, 2 out of 14 --or 14,29% of the group and 10% of total response. IDs **21**, **33**, **72**, **125**, **145**, **194** responded from the 10-19 group, 6 out of 54 --or 11,11% of the group and 30% of total response. User ID **158** responded with an average estimate of 4,4 for the site. **158** accessed only 5 web pages, and therefore the ID is not on the map for investigation. Only visitors of the two most active groups wrote constructive and critical commentaries with their evaluations, **47**, **21**, **85**, **120**, **131**, **93**, **153** and **204.** Evaluation was made on the scale from 1-5. The figure 0 is a bypass. [**Appendix I**]

Partaking visitors from the most active group (150+), consist of four scholars and one university student. Two scholars (**120**, **131**) and the student **(21**) research medieval Nordic literature, the scholars located in Iceland and the student for the most part in Germany. One scholar (**47**), from Australia can be assumed to research either medieval Nordic literature or Old Icelandic and comparative linguistics because **47** evaluates the site as a resource for those fields of study and abstains from evaluating it as resource in other areas. **85** is an Icelandic scholar, with no further identity. Their average use of the site is 273 web pages, in the range 205-558.

[10] A total of 117 displays of image versions for print were detected. Visitors may have printed out from the *full view* or *thumbnail view* displays. This was not recorded.

Partaking visitors from the 50-149 group are two scholars **(153, 204)** and one Australian university student (**93**). The student **93** is not further identified but **153** researches medieval Nordic literature and **204** researches in the field of Old Icelandic and comparative linguistics. Their average use of the site is 115, in the range 76-135.

Appendices I & J are referred to for information about other partaking visitors in the questionnaire, their user histories and response. It is however noteworthy, in the detailed account of partaking visitors from the two most active groups above, that 3 of them are located in Iceland, 2 in Germany, 2 in Australia and 1 in Sweden. Only one location (Iceland) offers direct access to the physical items. From the less active groups, 1 partaking visitor is located in Iceland and 3 somewhere in the USA. The others are located in Australia, Austria, Switzerland, Sweden, Norway and Canada

There is not enough data to distinguish between groups based on activity or on user status in conjunction with the response to the questionnaire. Only two patterns in the survey are noteworthy.

a) There seems to be a general agreement about the effectiveness of search and browse functions and help information, over three activity groups. The first column of fig.13 is for evaluating the functions themselves and the second for evaluating help information.
b) There is some diversity in the evaluation of effectiveness of image use and catalogue data over all activity groups. The first and second figure columns in fig.14 are for image use evaluation (for 1: research and 2: teaching) and the third column for evaluating catalogue data.

150+	3	3
	3	0
	3	3
	0	0
	3	3
50-149	3	2
	3	3
	3	3
30-49	4	3
	4	2
	3	4

(Fig. 13)

150+	3	3	3
	5	0	4
	4	4	5
	0	0	0
	5	5	4
50-149	3	3	3
	5	3	3
	4	4	4
30-49	5	2	3
	5	5	4
	5	5	5
20-29	4	0	0
	5	4	4
10-19	0	0	0
	0	3	4
	5	5	4
	3	3	3
	4	0	4
	4	4	4

(Fig. 14)

This diversity is less obvious though amongst those who belong to the same field of study, here medieval Nordic literature. They seem to agree most of them (Fig.15), on the usefulness of images for research (first column) and they are somewhat in agreement on the usefulness of catalogue information (third column). However, the evaluation of images for teaching (the second column) varies in grade from 2 (not very effective) to 5 (very highly effective). **153** made the remark that most students never get further then mastering normalized texts.

21 - (1)	3	3	3
120 - (1)	0	0	0
131 - (1)	5	5	4
153 - (1)	5	3	3
87 - (1)	5	2	3
72 - (1)	5	5	4

(Fig. 15)

Interface structure – design issues

Visitors from the two most active groups and that took part in the survey, wrote important critical commentaries. They fall into 3 categories and a few examples are introduced below.

1. Navigational structure, visual clarity
2. Catalogue data (detail of catalogue and conventions)
3. Download time of images

21, **153** and **204** complain about the depth of the navigational structure and talk about *backing in and out the tree*. When their routes are taken, according to their histories, it becomes evident that within a single session they back out to the front page from a full image view or a thumbnail view, in order to start a new browse. This can mean backing 3 or 4 steps when there is a different way to do this in 2 steps. At a closer examination it is concluded that a lack of clarity in the interface design causes some users to choose the longer route. IDs **447**, **208**, **150**, **131** for example do not have this problem but it might be wholly coincidental that they arrive at the shortest route, or it means that they are more experienced web users. This happens on 2-3 occasions for ID **12** and **60**.

Id **47** mentions a problem navigating the site, i.e. a problem finding routes. **47** is certain that the visual appearance of the site could be confusing to users because it is not immediately clear what are live hyperlinks.

ID **85** complains also about the navigational structure and suggests that a living human is needed to show what can be found and how. When the route of **85** is taken according to user history, it can be interpreted that **85** does not know what to expect of the site, i.e how this type of specialized source material is organized.

120 and **131,** both in the field of medieval Nordic literature and located in Iceland discuss the catalogue data. **120**'s concern is that a lot of more recent research information about some of the older manuscripts from the AM Institute is missing from this new digital catalogue. This comment is important, because it must be a continuing concern for the Sagnanet design team, to what extent manuscripts should be catalogued. The catalogue is extremely detailed as it is, including not only detailed structural information about the objects them selves, but also including many comments about the item's content and state of condition, which is according to catalogue conventions. However, it is clear that this is a continuing project.

131's concern is regarding the appearance of manuscript catalogue records and how they *look* compared to tables of content w/ pagination that give direct access to images. Conventional manuscript catalogues are designed to give as complete an *image* of the

manuscript as possible, in order for researchers not to have to touch an item with their hands, unless absolutely necessary. There is also a convention of how manuscripts are divided in to *parts* in the catalogues when they are written over periods of time and the parts are always numbered with roman numerals. Some of this was not accurately copied into the digital and interactive environment, but everything **131** pointed out can be fixed.

ID **93**, an Australian university student, is the only visitor who mentions delay on the Internet and claims that an otherwise wonderful access to images suffers the download time. It is recognized that Sagnanet images are high quality and often large and that the Sagnanet website will only be useful on the long run to visitors with high-speed connections to the Internet. This may explain the early on withdrawal of some visitors, for example from Russia.

- **Discussion**

This study took off with basic but broad questions, such as what is the geographical scope of users, how much material is fetched from the site compared to user hands-on demand at the local collections, what material is viewed and how, what do user histories look like, how does free-text search compare with browsing options, and how are images viewed?

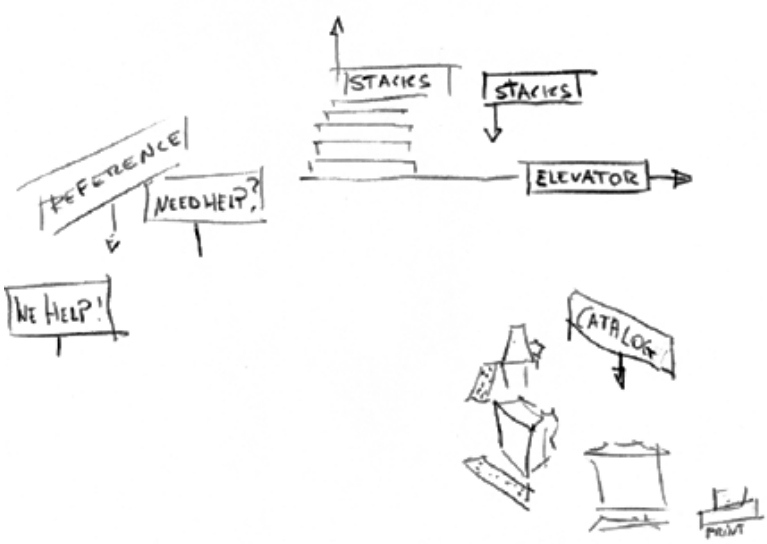
We have shown that there is increase in demand that can only be explained by the on-line and instant availability, but we have not shown that it will necessarily *better protect* the physical items from direct handling, which was one of the aims of the project. We still do not know if the effect of providing this service will result in more frequent visits to the collections to follow up on preliminary remote studies conducted in the virtual environment.

We have shown that combining three physical collections in one virtual environment is of direct interest to users and has great practical value in comparative studies. This feature was celebrated by a number of visitors and we see the outcome as one of the bigger successes of the project. We have furthermore shown that users generally have no problem using the site, however it was surprising how popular the browsing options are compared to free-text search. We have not shown precisely why this is the case, but the reasons we find most likely have to do with uncertainties of spelling and special characters and an obvious learning phase for new users, some that may not have high level computer literacy. We touched up on the interface design to meet fair criticisms of the appearance and the navigational structure, and we hope that the site has improved to user satisfaction.

Examining individual user histories was invaluable for better understanding of the usage and to substantiate (or invalidate) user responses, in particular letters of criticism and

free text comments with the questionnaire. One interesting issue remains though and kept cropping up throughout all our direct exchange with expert users.

The organization of the on-line material simulates the physical collections as much as that is possible. One collection is quite chaotic (NULI) another has a long history of organising and managing its manuscripts (AM). These two provide service of relevant expertise in the material content to their visitors and the on-line repository obviously does not. It is a feature impossible to simulate and even if the FIC collection of rare books and manuscripts does not offer such service locally, the FIC collection is a recognised social space where scholars are funded to meet, stay and study.



If we were to draw an analogy between Sagnanet and the most general university and national library setting, the limitations are already clear. Visitors use computer catalogues to locate items, either remotely or on site. Once they have located what they are looking for, or they think is what they need, they enter the physical space, *the stacks*, and travel so-and-so many floors up and down. In theory, this ought to work un-problematically and the only *real problem* is that the physical items are not all instantly available on-line in digital format and the wandering through stacks a forgotten history. But this is nowhere near to how a search for a book takes place at all times and particularly not a search for rare, ancient and highly specialised materials. One example of what often happens in the stacks is that the item is not present but the visitor can wander and look through so-called *nearby titles*, come across related materials and sometimes feel lucky to unexpectedly find a good source. This can be simulated in a design of a digital repository and it is very well done in Sagnanet by providing effective browse and search functions.

Another common scenario is to get back to the front desks and visit the persons sitting at the reference desk, presumably experts at handling and using information resources, -- *reference librarians*. There is no obvious way to simulate this feature and if the need was for live interaction with expertise in the *material content*, which is the case with Sagnanet, this is entirely impossible. Many of the items available in Sagnanet are clearly defined and

meaningful objects in close circles of experts. Some of them have accumulated an incredible knowledge of the collections' holdings. Furthermore, centuries of history support the culture as a whole in the way studies of the materials are conducted. It is precisely this recognition of libraries' special collections as social spaces where experts meet, discuss and examine, that draws obvious limits to on-line usability and perhaps can guide future designs of electronic access.