BibHit

A New Internet Search Tool Created by Danish Libraries

by [Karen Williams](#)

The article describes a new service, BibHit, provided by Danish libraries. The service was brought on by a wish to help library users become better acquainted with the Internet and at the same time add high quality Internet resources to the material made available by the library. The article describes how the BibHit model works, the criteria used for selecting valuable Internet resources and how people can participate in the project.

Contents

How do you as a librarian and mediator of knowledge respond to the abundance of information available on the Internet? Do you choose to ignore the massive flow of data because it is too enormous a job to handle the more than 1 billion pages on the Web? Or do you choose to take up the challenge and try to make access to the sea of knowledge as easy as possible? For example by developing special tools which both librarians and borrowers can benefit from?

Many libraries and librarians all over the world have been faced with this question for a number of years now and different solutions and ideas have come about in recent years. The following article is an introduction to one of the ways a group of Danish libraries have approached their role as mediators between their borrowers and the Internet.[1]

(To the top)

1. BibHit - A New Search Tool

BibHit is a new high quality search tool developed for people using the Danish libraries. By entering the Internet address http://www.bibhit.dk/ into your browser you have access to an effective search tool which can be used to find Internet resources which have undergone a high quality check and been catalogued in the database of BibHit. If you have the need to find more than Internet resources, you can extend your search to include the library's own collection of books, CD's, articles etc. As a result of this search you are given a complete list of all the material the library has to offer you in regards to your topic of interest.

To put it another way - in case the library cannot present a book or an article covering a specific topic, it will be possible to present Internet resources of high quality instead. In theory, Internet resources are always available. In this way the library will always be able to help its borrowers with useful material about a desired topic.

(To the top)

2. The BibHit Model

The name "BibHit" is derived from the Danish word for library (bibliotek) and the English word "hit" which relates to the results you get when having performed a search in any of the many search engines on the Internet. It also refers to hitting the bull's eye in the sense that all the hits you get when searching BibHit's database are of high quality.

BibHit was developed by a group of Danish libraries; more specifically three public

libraries, Hobro Public Library, the Main Public Library of Aarhus and Silkeborg Public Library, and one research library, the Library of the Aarhus School of Business, between 1998 and 2000[2]. In June 2000 the research library withdrew from the project, thus leaving the three public libraries to complete the project. During the mentioned time period the project received financial support from the Danish Ministry of Culture through the National Library Agency's special pool for project developments – "Udviklingspuljen for folke- og skolebiblioteker"[3].

During the past year from June 1999 till June 2000 a model for collecting, evaluating and cataloguing carefully chosen Internet resources has been developed. As a result of this it is possible to search directly in a database full of references to a constantly growing number of full text documents found on the Internet by using BibHit's search interface at http://www.bibhit.dk/.

By making use of a special program, which uses the Z39.50 protocol[4], it is possible to search the library's own database which contains information about its collection of books, magazines, CD's, CD-ROM's etc. This is done parallel with the search in BibHit's database. The results are shown in one user interface in the form of a long list, which gives you an overview of the material available. The only thing needed in order to make use of BibHit is access to the Internet. This means you do not have to be physically at the library if you want to use BibHit - you can also access it from your own computer.

3. Internet Resources

What kind of Internet resources are you able to find in BibHit's database? First and foremost the database contains references to full text documents, which means that you will not find the opening page to the individual websites. An example would be the American newspaper The New York Times[5], whose opening page is not of interest to BibHit, whereas the articles which are to be found within the website are obvious candidates for inclusion in BibHit's database. They can be included because they fulfil the first selection criteria by being full text documents.

In other words, the pages catalogued are often found by exploring each individual website in depth. The reason for choosing full text documents is so the borrower only has to click twice with the mouse in order to reach the desired information after having searched BibHit and been given the list of hits. As a result of this rule the borrower is taken directly to the desired information and need not search a given website herself.

However, as in most other cases there is an exception to this rule. If the opening page to a website gives a clear idea of what can be found at it, and it covers a specific topic, the front page will be catalogued. An example which illustrates this rule is the official website for The Empire State Building.[6] It was judged to cover a narrowly defined topic – The Empire State Building – and the pages found at the website offered information specifically relating to the building.

4. Languages

The goal is to find Internet resources in eight different languages - Danish, English, French, German, Italian, Norwegian, Spanish, and Swedish. However, as a starting point the focus is placed on Internet resources available in the Scandinavian languages as well as English. The text reading level begins at a level comparable to the study level found among students finishing the last few years in the Danish Folkeskolen (ages 16-17) and goes to the Danish university study level. In regards to contents, any topic is of relevance, and the information can be of general interest or so specialized that it can be used for studies and research at the university level. Thus, BibHit is a search tool, which is particularly well suited for the librarians at the Danish libraries as well as the grown-up borrowers.

([To the top](#))

5. Coverage

Atypical of the Danish library system, an elaborate system of subject specialists has not been found necessary in connection with BibHit. This means that an Internet resource does not have to undergo an extensive and time consuming evaluation by a specialist before it can be added to the database. Within a few minutes after a resource has been found, it can be catalogued and found again in BibHit. As a result of this process, a person working as a cataloguer for BibHit will find herself cataloguing Internet resources covering a wide range of topics - a situation resembling the one most of the cataloguers at the libraries find themselves in already.

However, in order to secure a reasonable coverage of topics in the database, a decision has been made to make a list of topics, which it will be possible to find Internet resources about in BibHit. The inspiration for the list and its structure will be found in the official Danish library classification system DK5[7], the traditional Internet directories such as dmoz – open directory project[8] - as well as other Danish projects such as FNG.[9] The idea is that individuals working for BibHit will be able to focus their work on a specific field of interest, which will make the job more enjoyable and also guarantee the quality of the Internet resources catalogued in BibHit. However, the distribution of topics to a number of editors is not aimed at stopping participants from cataloguing resources, which are not related to their chosen topic. Thus, the database will eventually contain references to high quality Internet resources covering a huge variety of topics, but the system ensuring the coverage will not be slow and rigid to work with since it does not require that only certain individuals can catalogue resources about certain topics.

([To the top](#))

6. How to Catalogue a Resource

In BibHit, the cataloguing process has been made as uncomplicated and user friendly as possible. A number of helpful texts and examples as well as special online user

manuals have been made in order to help the cataloguers of BibHit[10]. The cataloguers can choose between two different templates for cataloguing Internet resources - a simple one and an advanced one. Both templates are based on the Dublin Core standard[11] - an international standard for making metadata. Among other things, the standard is used internationally in connection with libraries cataloguing Internet resources for various databases. Metadata is often displayed as part of the HTML markup, and can then be viewed in the web page source. It reveals certain information about the availability of the resource, its origins and its contents as well as its format. In regards to the time spent on the actual cataloguing process, the goal is that it should only take the experienced BibHit cataloguer between five to ten minutes to create a reference to a chosen Internet resource and save it in the database.

(To the top)

7. Collecting Resources

The process of collecting the Internet resources consists of three different methods. The first procedure makes use of the group of professional and experienced collectors of links working for BibHit. By making use of the traditional search engines and netguides already on the Internet, they carry out goal-oriented searches. The resources they find they will either add to the database themselves or send by e-mail as a suggestion to another group of people involved in the work, i.e. the cataloguers who do not search for resources themselves. The second method involves a so-called harvester robot, Combine Harvester.[12] This is basically a special piece of software, which can be programmed to collect Internet resources by taking its starting point in a specific URL. Once the harvesting is over and the resources have been screened and duplicates removed automatically, the remaining resources will be evaluated and added to BibHit's database if they pass the test. The third and last method makes use of the borrowers at the libraries. They are given the option to suggest Internet resources, which they think might be relevant to include in BibHit. However, one thing all three methods have in common is the fact that a person has to evaluate the Internet resources before they are added to the database.

(To the top)

8. Maintenance

Internet resources have always been famous for being unreliable - a characteristic that could cause lop-sidedness in a project like BibHit if the database was not kept up to date and the URL's checked regularly. However, BibHit has a special computer programme, which automatically checks if the catalogued resources are still available. At present the check is carried out twice a week and at two different times of the day. If the system has been unable to reach a page, an e-mail will automatically be sent to the person who catalogued the Internet resource. Then it is up to the cataloguer to investigate what happened to the page. In other words, you are responsible for the resources you have catalogued but at the same time you are not asked to spend time weekly or monthly manually checking if a page has disappeared from the Internet. The check is done automatically.

9. Background and Motivation

BibHit emerged from a wish to create a way of presenting Internet resources to library users and making the librarians think of the Internet as a good source of information. In most cases the Internet is not the source of information most librarians look to when they try to answer the questions posed by the borrowers. More often they will choose the traditional sources, such as books, articles etc., even though they may have the feeling that useful information can also be retrieved from the Internet. The motivation behind this choice is grounded in an impression of the Internet as being vast and difficult to navigate. This, combined with a lack of knowledge about the traditional search engines, their development and use, makes you even less inclined to take the plunge into the Internet. With the creation of BibHit, an easy entrance to the Internet has been made which relieves you from having to search the entire Internet for pages of high quality.

BibHit also works as an incredibly effective and time saving tool for librarians and borrowers alike. With BibHit the valuable Internet resources, which have already been found and catalogued in the database, do not have to be found over and over again. As long as the people participating in BibHit remember to catalogue the useful resources right away, and people in general remember to use BibHit, everybody can save time and energy in their working processes.

10. Another Year

The project applied to "Udviklingspuljen for folke- og skole-biblioteker" for another sum of money to cover the expenses for yet another year. At the beginning of July the message was clear: BibHit had been granted enough money to take the project to the end of 2000 and at the same time had been invited to apply for money to make BibHit a permanent library service. By mid November 2000 it will be clear if BibHit is going to change from being a project limited by time and money to a permanent search tool available to librarians and borrowers alike. However, if the necessary sum of money is not granted, the project will apply to another library fund for money in order to be able to finish the project in July 2001.

11. The Future

During the coming year, the focus will be placed on four different areas. First of all, the project would like to expand the number of libraries and individuals participating in the work. Secondly, the user interfaces already created will be made even more user friendly and more functions will be added. Thirdly, at the end of the second project year BibHit must be able to stand on its own feet, which means that the conditions for a permanent running of BibHit must have been decided upon. Finally, BibHit will be

presented to the users of the libraries. This has not yet been done due to the fact that BibHit's database of references to Internet resources must be of a certain size before it can meet the needs of the borrowers concerning their desire for coverage of a wide range of topics. By expanding the group of participants it will be possible to meet these requirements.

12. Participants

BibHit welcomes new participants to the project - whether they are libraries or individuals. One of the main requirements is that in order to be part of the project you must have a connection to the library world. However, that does not mean you have to be a librarian and a subject specialist in order to contribute to BibHit. People with that particular background are, of course, welcome to do so. However, the founding libraries quickly realized that in order for the database to gain a reasonable volume and have an acceptable amount of topics within a few years it would be sensible to let library assistants, library cataloguers and students at The Danish School of Librarianship become involved in the project.

BibHit has only recently started its campaign of inviting new participants to join the project, but the number of cataloguers is slowly increasing.[13] The goal is to have at least 200 people working at least one hour a week collecting and cataloguing new resources as well as checking old ones that have either moved or disappeared completely from the Internet.

13. More information about BibHit

If you would like to know more about BibHit it is possible to find more information on the web site for the project: http://www.bib.dk/info.htm. You can find the official description of the project, read the applications for state grants, study the list of FAQ's and the collection of links, find information about the libraries, read the online manuals and a number of other documents telling you the story of the project and its development, and get a list of the companies and people involved in the project. You can also contact the project manager for BibHit, Karen Williams, by e-mail info@bibhit.dk for more information and details.

About the Author

Karen Williams holds a Masters Degree in English and Danish from the University of Aarhus, Denmark and currently works as project manager for the library and Internet project BibHit (http://www.bibhit.dk/). She also holds a part time job as co-editor of Systime EnglishWeb - a collection of Internet resources to be used in connection with the teaching of English at High School level, which is published on the Internet by the

Danish publisher Systime (http://www.systime.dk/).

Notes

1. For other examples of how Danish libraries have managed to make use of the Internet and present the information found on it to their borrowers see for example Folkebibliotekernes Netguide (FNG) http://www.fng.dk, Biblioteksvagten http://www.biblioteksvagten.dk/ and Mimers Brønd http://www.silkeborg-bibliotek.dk/mimer/   (Back to the text)

2. All four libraries have web sites - Hobro Public Library http://home1.inet.tele.dk/hobrobib/ , the Main Public Library of Aarhus http://www.aakb.dk/ , Silkeborg Public Library http://www.silkeborg-bibliotek.dk and the Library of the Aarhus School of Business http://www.lib.hha.dk/ (Back to the text)

3. At the homepage for Biblioteksstyrelsen http://www.bs.dk/ you can find more information about Udviklingspuljen for folke- og skolebiblioteker.   (Back to the text)

4. See http://lcweb.loc.gov/z3950/agency/   (Back to the text)

5. See http://www.nytimes.com/   (Back to the text)

6. See http://www.esbnyc.com/   (Back to the text)

7. See http://www.kb.dk/kb/dept/nbo/da/samlinger/dk5/oversigt/   (Back to the text)

8. See http://dmoz.org/   (Back to the text)

9. See http://www.fng.dk/fng/   (Back to the text)

10. The cataloguers in BibHit make use of the "Materialevalgsmanual" http://www.bibhit.dk/info/mat_man.htm (Selection Criteria) and the "Katalogiseringsmanual" http://www.bibhit.dk/info/kat_man.htm (How to Catalogue a Resource)   (Back to the text)

11. See http://dublincore.org/   (Back to the text)

12. See http://www.lub.lu.se/combine/   (Back to the text)

13. For at quick estimation of the number of people involved in BibHit, take a look at http://www.bibhit.dk/info/katlist.htm   (Back to the text)

Return to Human IT 4/2000