

Multilingual Web Services of Data Archives: Possibilities and Pitfalls

Our article discusses the kinds of issues data archives and other data providers need to consider when setting up multilingual websites. The article, based on a presentation at the IASSIST 2007 conference (Jääskeläinen and Alaterä 2007), draws from our experience at the Finnish Social Science Data Archive (FSD), which provides web services in three different languages, and from the examination of other European data archive websites.

First things first

When planning a multilingual website, we recommend that the purpose and scope of the site be discussed and decided first. Website developers need to ask themselves the following questions: why are we making the website multilingual, what kind of functions or information do we wish to offer, who will use the service, and what are the users' needs. Having web services in multiple languages means more work, so they must also consider how much money and staff resources can be allocated. Organisations often have to balance between maintaining an adequate level of services and having too much information to update.

Clear decisions regarding goals and content should be the first step in the planning process. For example, the FSD provides web services in Finnish, English and Swedish, but has different goals for each. The language of the main site is, of course, Finnish. Since five percent of Finns count Swedish as their mother tongue, we also offer rudimentary web services (such as information on how to order and deposit data) in Swedish. Our English website is fairly comprehensive in order to provide services for all non-Finnish speakers. One of our main goals is to provide information on data, therefore we have study descriptions in English for all archived quantitative datasets. Roughly 80% of the web pages in Finnish have a corresponding page in English, though the content is not identical.

Priorities for data archives

When we look at data-rich, multilingual websites from the user's point of view, we must ask ourselves what are the main reasons for users to visit such sites? It is a fairly safe bet to assume that most users will be looking for data and information on data (Hansen and Richardson 2006).

*by Taina Jääskeläinen and
Tuomas J. Alaterä**

However, in our experience browsing the multilingual websites of European data archives (mainly the English versions), it sometimes seemed easier to find information on the archive than on data!

We found, quite often, that links from an English page led, without any prior warning, to a web page in another

language, often to the data search or data catalogue page in the dominant language. This can be considered a major flaw in web design as well as in usability. An arrangement like this can work if many users have at least a passing knowledge of the dominant language; however, when it is reasonable to assume that foreign users do not understand the dominant language at all (as is the case with Finnish), some additional information is needed.

We recommend having a declaration of content that provides details on the web services that are available in each language. This declaration will prevent users from having to browse through the whole site in order to find out what services and information are available in a particular language. The declaration need not be long but it must be to the point. A couple of well-placed sentences will most likely be sufficient.

Since many users of multilingual websites are looking for data, the declaration of content for the non-dominant language pages should answer the following questions:

1. Is it possible to search data in that language?
2. Are there study descriptions in that language?
3. If there are no study descriptions, who is the contact person for information on the data?

Even without the declaration, users will appreciate finding answers to these questions quickly and easily. If registration is required at any point, providing information on who can register is a good idea. If it is available for all, say so. What may be self-evident to those who are setting up a website is often much less so to users.

Let us assume that the user has found data that are relevant to him/her. Now he/she needs to know how to proceed. A

functional website provides adequate and easily-accessible information on how to order data, for example, by providing a link from the study description page to ordering information and order forms. Users need to know whether (and how much) it may cost to order data, whether the archive can provide translation if requested, and how much the translation will cost. If the archive does not provide data translation, stating this will make the situation clear to the user without further enquiries.

If the data archive does not have the resources to provide study descriptions in the non-dominant language(s), one alternative is to translate the data search interface, including DDI field names. This may be useful for users who have a passing knowledge of the dominant language. If they have access to the search interface in their stronger language, they may be able to use dictionaries or multilingual thesauri to find search terms for study descriptions in the dominant language.

We also recommend providing detailed contact information. That is, more than just one email address and telephone number for the whole archive. Users appreciate knowing who does what and who they can contact regarding a particular issue.

It is worthwhile to carry out a user test before launching a multilingual website. A simple and quick one may be sufficient. Recruit a couple of people not familiar with the creation process of the site and ask them to spend up to an hour trying to find, for example, information on data on a particular theme. Have a staff member sit beside the tester and take note of his/her comments. It is crucial that testers try to use the services offered on the website and not simply review how the site looks. That being said, a critical look at colour etc. may also be useful. In all probability, the hours spent on this type of simple testing will turn out to be very useful for the eventual functionality of the site (Alaterä 2000).

About translation

When considering translation of a website into another language, it is important to remember the differences that may exist in a particular language that is spoken in more than one country. For example, at the FSD we had to consider whether we were designing our web pages in Swedish for users living in Finland who count Swedish as their mother tongue, or people living in Sweden, or both.

People living in another country may not be familiar with your society and systems. This fact is particularly relevant when dealing with abstracts in study descriptions, as some terms related to health care, taxation, educational systems, etc. may not be understandable to users from other countries, even if they speak the language. It may be better to globalise some terms instead of merely translating the abstract as it is. The procedures for ordering data may

also be different for researchers working in your country and for researchers working in another country. Ordering information should clearly reflect this.

The people who are paying most detailed attention to the content of a multilingual website are the translators. Therefore, if translation is contracted out, the translators should be well-informed as to the goals and target audience for the site. This will make it easier for them to decide when the content needs to be adapted to fulfil the goals (Nielsen 1993, 242–245). Even if the person doing the translation is a staff member, process writing and teamwork is always useful. We suggest pre-writing, consultations with other staff members to discuss the result followed by revision, user tests, and further revision. Many heads are better than one.

Easy navigation

Clarity of navigation is a must, whatever the language. It is a well-known fact that Internet users are impatient. It has been said that “if a man from Mars doesn’t figure out your navigation in four seconds, your web page sucks” (Flanders). Even though somewhat provocatively expressed, these are wise words and well worth keeping in mind! (If you would like to see other web design tips like this, go to <http://www.webpagesthatsuck.com/>.)

Two main pitfalls in navigation are making the navigation meet the needs of the organisation rather than those of the user and assuming that users know more of the organisation and its services than they in fact do. If the website does not provide one specific link to the data catalogue, but has several links to different types of surveys (for example, different survey series), it may confuse users. They will have to go to the page of each different type of survey and try to figure out how they are categorized and what the content is. Navigation problems like these are easily revealed by a simple user test. Again, some kind of declaration of content will be beneficial (Hansen and Richardson 2006).

Regarding links between language versions, frequently used practices are the best (W3C Working Group 2007). We recommend putting language links where users expect to find them and always in the same place on every page. If the language links are flag-based, adding or replacing these with text (English, español, deutsch) is strongly recommended. People do not remember flags, and, in the case where one language is spoken in several countries – which flag would you choose?

When providing links from one language version to another, the optimal solution is to provide a direct link to the corresponding page in another language. However, this is often difficult or impossible since the content of different language versions differ. It also means more work. Linking to the home page in the other language is the most common

solution, which again stresses the importance of a clear and understandable navigation design. Users will be keen to get back to “that one important” page instead having to test their search skills trying to figure out where they might find it. This task is made easier if bread crumbs are used on each page and, if possible, each language version should be given similar structure (Nielsen 2000, 325–331).

Web design

We recommend that plenty of time be dedicated to discussing web design. Concentrate on achieving your desired goals. Web design should be based on the goals and target audience chosen for each language version. What functionalities might target users expect (Felke-Morris 2006)?

In all web design it is important to remember that Internet users are likely to be impatient. Avoid splashes and unnecessary animation, especially on the index page of your website. Has anyone ever seen an animation that they would like to see again, again and again? It is advisable to think twice before choosing to rely solely on for example Flash techniques or additional plug-ins. Difficult or non-conventional navigation (which means that users get lost and frequently have to return to the main page) combined with time-consuming animation on the main page usually results in very frustrated users. This type of design does not work for information-rich websites where it is crucial to be able to find a particular piece of information without too much effort. This is why popular web services like Google rely on a simple and fast-loading front page.

Applying some kind of a template system reduces the burden of running websites in multiple languages. When changes are needed in navigation or in other fairly constant elements of the site, templates make life much easier. Fonts, colours, titles, navigation links and so forth are controlled by one style sheet and one design template, meaning that only one alteration per language is sufficient. Using templates also tends to have the effect of pushing the web design toward an easily-maintained form, which is good for everyone.

It is a good practise to put navigational elements (search interface, database field names, etc.) into language packs, and to keep them as separate as possible from the text content of the site. This is preferable to coding these elements into the source code directly. With language packs, it is easier to translate functionalities like navigation, search interface etc., and to make changes in all language versions simultaneously or even to add completely new language versions.

When a website is up and running it should not be left on its own. Checking the log files periodically gives valuable information on where users are coming from, how long they spend on the site, if there are pages that are more

popular than others, and which pages (if any) tend to drive people away to information sources outside your site (Rasmussen 2007). Analysing search logs helps to find out how and what people are searching for, and whether they are looking for information that does not exist on your site, or whether they are looking for existing information but in a wrong place. Log file information can be used to improve your web service.

Conclusions

The main point of our discussion is that multilingual websites should be designed as if they were separate websites developed in separate languages, rather than multiple translations of a single language site. What we are talking about here is creating new websites in other languages. What the content and actual technical solutions are depend very much on the goals set for each site. Decide on your goals, be clear about what services you are providing, and never lose your users to difficult navigation.

You might ask do we at the FSD have optimal multilingual web services? The answer is probably not, but we did recently make changes to our multilingual websites in order to follow our own recommendations. There is always room for improvement!

* This article was originally presented at the IASSIST 2007 conference. The authors Taina Jääskeläinen and Tuomas J. Alaterä work at The Finnish Social Science Data Archive, 33014 University of Tampere, Finland. Contact: taina.jaaskelainen@uta.fi

References

- Alaterä, T. J. 2000. "WWW-suunnittelun periaatteita" (lecture presented at the Institute for Extension Studies at the University of Tampere).
- Felke-Morris, T. 2006. *Web Development & Design Foundations With XHTML*. 3rd ed. Boston: Addison Wesley. Best Practices Checklist also available online at <http://terrymorris.net/bestpractices/index.htm>.
- Hansen, S. E. and M. Richardson. 2006. Evaluation of web sites: What works and what doesn't. Paper presented at the IASSIST 2006 Conference, Ann Arbor. Also available online at http://www.iassistdata.org/conferences/2006/presentations/C2_HansenRichardson.ppt.
- Jääskeläinen, T. and T. J. Alaterä. 2007. Multilingual web services: Possibilities and pitfalls. Paper presented at the IASSIST 2007 Conference, Montreal. Also available online at http://www.edrs.mcgill.ca/IASSIST2007/C3_FSD.ppt.
- Nielsen, J. 1993. *Usability Engineering*. San Diego: Morgan Kaufmann.
- Nielsen, J. 2000. *WWW-suunnittelu*. Jyväskylä: IT Press.

———. *Designing Web Usability*. Indianapolis: New Riders Publishing.

Rasmussen, K. B. 2007. F2: Data beyond numbers: Using data creatively for research. Paper presented at the IASSIST 2007 Conference, Montreal. Also available online at [http://www.edrs.mcgill.ca/IASSIST2007/presentations/F2\(1\).ppt](http://www.edrs.mcgill.ca/IASSIST2007/presentations/F2(1).ppt).

W3C Working Group. 2007. Internationalization Best Practices: Specifying Language in XHTML & HTML Content. World Wide Web Consortium. <http://www.w3.org/TR/i18n-html-tech-lang/>.

Flanders, V. Does your web site suck? Flanders Enterprises. <http://www.webpagesthatsuck.com/does-my-web-site-suck/does-my-web-site-suck-intro.html>.