Enabling Resource Selection Based on Written English and Intellectual Competencies

Ayako Morozumi[1], Liddy Nevile[2], and Shigeo Sugimoto[1]

[1] University of Tsukuba, 1-2, Kasuga,
Tsukuba City, Ibaraki, Japan
{moro,sugimoto}@slis.tsukuba.ac.jp
[2] La Trobe University, Kingsbury Drive,
Bundoora, Australia
liddy@sunriseresearch.org

Abstract. A growing number of people are using the Web to access English-language resources, among other things. In Asian countries, for example, many people want access to English texts. Many Asians are not as competent reading English as they may be in the intellectual content of their domain. The problem of accessibility to English texts is significant simply because of the number of people involved. The problems for second language English readers are similar to those for many dyslexic first language readers. We propose a descriptive model that supports adaptability of texts for the benefit of such people based on FRBR and AccessForAll standards.

Keywords: FRBR, AccessForAll, resource descovery, selection, adaptation, accessibility, English second language, ESL, dyslexia.

1 Introduction

There are at least three major groups of readers with language-skill problems who want access to intellectually stimulating and specialist English texts:

- people with domain expertise who lacking sufficient English reading skills to access the English literature in their field of interest;
- people with domain expertise who need translations of English literature, and
- people with dyslexia.

Although there are texts in many languages, there is sufficient interest in English literature for it to be the focus of this paper. Previously we described a metadata schema model for users seeking appropriately 'accessible' resources [1]. In this paper, we focus on the selecting a resource depending both on the knowledge level of the user and the (sometimes second-language) reading skills of the reader.

We consider the problem for second-language readers, translators (particularly automated ones) and people with dyslexia to be similar: in all cases it is important to have plain English without distracting or confusing metaphors, or complicated language constructions such as the subjunctive mood or passive voice.

D.H.-L. Goh et al. (Eds.): ICADL 2007, LNCS 4822, pp. 127–130, 2007.

2 Research Base

We follow the AccessForAll model of accessibility [2] in which a user specifies their needs and preferences and a resource is discovered and, where necessary modified, to match these requirements. Accessibility, in this sense, means that the display, control and content of the resource is suited to the user, regardless of any disabilities they may have or circumstances causing a lack of access. In this paper, the lack of access of concern is to content that is originally English.

We consider that unless resources are suitably described for both their content level and their reading level and these two properties are related, users will not have be able to find suitable resources for their personal use. Further, we are concerned that the reading level of a text in a second-language is not the same as the reading level for a first-language reader, and that this should be described differently.

2.1 Functional Requirements for Bibliographic Records

Although the Functional Requirements for Bibliographic Records (FRBR) [3] were originally developed for books, they are increasingly proving useful for digital resources. In this paper, we want a resource that contains the intellectual content that is in the original *work* that resulted in an original *expression*. Having found the right *expression*, the user will need to discover a *manifestation* of that expression to access as a suitable item when delivered. These are the four levels of abstraction, called entities, in the FRBR model.

2.2 AccessForAll

The AccessForAll metadata model depends upon a description of a user's access needs and preferences being specified and matched by the characteristics of a resource, as described in the resource's metadata. This, of course, requires the availability of a matching service.

The innovation in the AccessForAll approach to accessibility is that the matching service can enable cumulative, distributed components to be combined to make a resource more accessible, even after the resource has been published in its original form. In some cases, this involves the de-construction of a resource into components and the re-construction including alternative, adapted or augmented components.

It is the just-in-time accessibility that is to be exploited in this paper. This means that a Web service, for instance SWAP [4], could render a resource accessible by providing a translation of it, or a plain English version.

The AccessForAll approach advocates description of both needs and preferences because for some people a need is crucial, and if not satisfied the resource will not be useful at all, while for others the stated need is a preference, and if not satisfied, may make for difficulties that will be tolerated by that user.

3 Metadata Standards and the Needs of the User

The purpose of our study is to provide useful models for resource characteristic description and possibly for complementary needs and preferences descriptions. What

is needed is a way of making it easy for users to have their discovery results matched to their needs and preferences. The following paragraphs discuss metadata standards for resource discovery and functional augmentation for resource access from the viewpoint of accessibility.

3.1 Metadata Standards

The AccessForAll specifications have been formalised by the International Standards Organization [5]. They provide for two descriptions: the description of the user's needs and preferences and the potentially matching properties of the resource. This model can be extended simply as it depends simply upon the presence of the descriptions and the matching service.

We consider that a user first needs to discover a resource based on the intellectual content of the resource. FRBR provides a standard framework for such a description and it has been partially or completely implemented in many available standards, such as Dublin Core, MARC-21, MODS, etc.

Having found the work and the expression, the problem is to find a suitable manifestation to deliver the right item. This means a manifestation that has suitable characteristics such as that it is in plain English, which can be easily and most accurately translated, read, or interacted with by a person with dyslexia.

DC Audience [6] is a standard term used to describe the class of person to whom the resource is aimed, or for whom it is considered suitable. The class of people is usually expressed as citizens, or Grade 4 children. AccessForAll avoids such judgments in favour of descriptions of properties of the resource so that individual users can assess the suitability of the resource for them as individuals. This is significant because users do not have the same needs and gross classifications can eliminate resources a particular user could enjoy.

In MARC-21, there are elements such as *Reading Grade Level* and *Interest Grade Level* [7]. We foresee no problem with using such established standard descriptions but note that it is not the author or publisher's target audience that should determine the access but rather the user's needs and preferences.

We are arguing therefore, for a standard way of expressing these in combination that can be accessed by AccessForAll services. The AccessForAll model provides a way of doing the matching and applications such as The Inclusive Learning Exchange (TILE [8]) already show how implementations might work.

3.2 Functions to Augment Accessibility

Web services are emerging that are capable of automating such processes. There are automatic translators. There are online services that offer immediate translation by humans. There are services that simplify texts in appropriate ways for some readers, transform them, we might say. SWAP is an example of such a service [4].

This paper does not call for a new model so much as an additional set of elements developed according to the established model. We argue that once the relevant characteristics are ascertained and standardized, determining the relevant values is not necessarily a human task. Reading levels have been determined automatically in the past. Potentially, second-language reading levels could similarly be determined.

'Clutter' that causes serious problems for dyslexic readers may be an example of what is relevant, as would be the use of metaphors, passive voice, subjunctive tenses, etc.

4 Conclusion

In this paper, we have drawn attention to what, we believe, is a significant problem in quantitative as well as qualitative terms. We adopt the position that the World Wide Web and associated technologies and practices have opened the way for wide international participation in intellectual endeavour but that currently there is unacceptable effective discrimination against non-English readers. We propose some requirements for consideration if this situation is to be averted. We argue that by describing the English expression within the text in appropriate, standard metadata, a significant benefit would be derived from a more accessible internationalized knowledge base.

References

1. Morozumi, A., Nevile, L., et al.: Using FRBR for the Selection and Adaptation of Accessible Resources, International Conference of Dublin Core and Metadata Applications. In: Proceedings of DCMI Conference 2006, Mexico, University of Colima (2006)
2. Nevile, L., Treviranus, J.: Interoperability for Individual Learner Centred Accessibility for Web-based Educational Systems. IEEE TCLT's Journal of Educational Technology & Society 9(4) (2006)
3. International Federation of Library Associations and Institutions: Functional Requirements for Bibliographic Records (FRBR), http://www.ifla.org/VII/s13/frbr/frbr.pdf
4. Semantic Web Accessibility Platform, http://www.ubaccess.com/swap.html
5. Heath, A., Nevile, L., Treviranus, J.: Individualized Adaptability and Accessibility in E-learning, Education and Training Parts 1, 2 and 3, http://jtc1sc36.org/doc/36N1139.pdf, http://jtc1sc36.org/doc/36N1140.pdf, http://jtc1sc36.org/doc/36N1141.pdf
6. Dublin Core Metadata Initiative, http://purl.org/dc/terms/
7. http://www.loc.gov/marc/bibliographic/ecbdnot1.html#mrcb521
8. The Inclusive Learning Exchange (TILE), http://www.barrierfree.ca/tile/index.htm