Managing electronic content for adaptation to the reader's profile: project MultiAbile for the inclusion of impaired e-Learners

Thimoty Barbieri[1], Antonio Bianchi[1], Licia Sbattella[1], Ferdinando Carella[2], Marco Ferra[2], Piergiovanni Santini[2]

[1]Dipartimento di Elettronica e Informazione, Politecnico di Milano, Italy
[2]Web Learn srl, Via Ronchi 39, 20133 Milano, Italy

thimoty.barbieri@polimi.it, abianchi@elet.polimi.it, licia.sbattella@polimi.it,
{f.carella | m.ferra | p.santini}@weblearn.it

Abstract

Project MultiAbile (http://www.multiabile.it) is an ongoing research and implementation effort to create a distance learning environment dedicated to the general public and to impaired people, offering a range of different channels and modalities to access the content and to study it. Some of the different modalities to access the content are: improved accessibility for screen readers; dynamic transcoding to synthesized speech over the telephone, with a dialogue-based navigation within the content; sound/tactile description of images (e.g. cartographic data) using force-feedback mice or pads. The content is also offered in two different textual versions, an original version and a re-processed version, in order to obtain high readability and understandability. The edited version is created in a way that only words from a basic dictionary are used (the Italian "Vocabolario di Base"), and the text is processed in order to obtain a high GULPEase index (an index of complexity in the structure of text). Users can describe their "profile" – whether they use a screen reader, or prefer to have a dialogue-based navigation, or if they prefer a simplified version of the text before confronting themselves with the original version. The user profile is used to dynamically adapt the content to the user, facilitating user access to it.
In order to encompass the variety of channels and modality with which the user can "read" content in the platform, we devised a metadata schema with which organize the elements of the provided content, and a creation and editing workflow to process the content before publication in the learning environment. The resulting content is also structured in a way to be compliant to SCORM standards, so it is viable to process for multichannel and multimodal access also pre-existing SCORM contents. In the paper we describe the general data structure devised within the project to accommodate the different access modalities to the content, and to allow the user to access and read it dynamically according to his/her profile. We explain how this structure impacts on the workflow of creating and editing of the content, or of repurposing and adapting pre-existing content in SCORM-compliant format. As a conclusion, we advocate that this workflow and content structure can pave the way to the offering of the same content to different channels and with different modalities (including emerging channels, such as Digital TV, and specific modalities, like for example automatic rendition to sign language), limiting the effort of repurposing the content by means of automatic transcoding and transformation algorithms.

1 Foreword and Motivations: Electronic Content Accessibility

The need to provide better accessibility to networked information is widely recognized and expressed in a number of initiatives and government regulations (W3C WAI, Section 508, etc.). However, most part of these guidelines focus on making more accessible the "navigation" or "publishing" environment (in the sense of the digital system through which the user can obtain the desired information, for example a web site through which the information is offered) more than facilitating the use and the comprehension of the electronic content within. Accessibility guidelines, in their current implementation, are mostly used as prosthetics, alongside other assistive technology devices (screen reading software, Braille displays, etc.), with web sites or publishing environments featuring an "accessible" portion alongside the "normal" part of the environment. While accessibility of the navigation/publishing environment is a necessary foundation, we advocate that much more attention should be devoted to the creation, aggregation and processing of electronic content, more than to the environment which is used to publish it to users. Improved strategies at content level allow to empower the user

offering him/her several alternatives to navigate and peruse the content itself, turning the way accessibility is currently implemented in electronic systems from a "prosthetics" notion to a truly multichannel and multimodal paradigm [2, 3, 12].

Focusing on how electronic content is structured, more than on the web site in which it "lives", allows to dynamically repurpose it with transcoding and transformation procedures over different communication channels and interaction modalities (e.g. display on a web page with AAA-conformance, read aloud with TTS over the phone [4], convert into Braille or in an accessible/readable PDF eBook, provide information via a Digital TV screen). This way, once the user has selected the content he/she wishes to access, he/she can select the preferred channel and plurality of modalities. In a scenario of content adaptability over the channel chosen by the user, it is crucial that the user can express, in a formalized way, the features which are required in accessing the electronic content. This information, collected and organized in a general model, the *user profile*, is the base to describe the interaction modalities between an electronic publishing system providing contents and its user, impaired or not.

In this paper, we present this approach, that we followed during the development of project "MultiAbile" (http://www.multiabile.it). We created a "publishing multimodal platform", which can deliver electronic contents in different forms, based on a simple user profile described by the user himself. The content was described using a simple DTD devised with these objectives in mind [12], and aggregated following SCORM standards to provide the functionalities of a multichannel, multimodal e-Learning platform which could offer improved accessibility to the teaching material for the impaired learner.

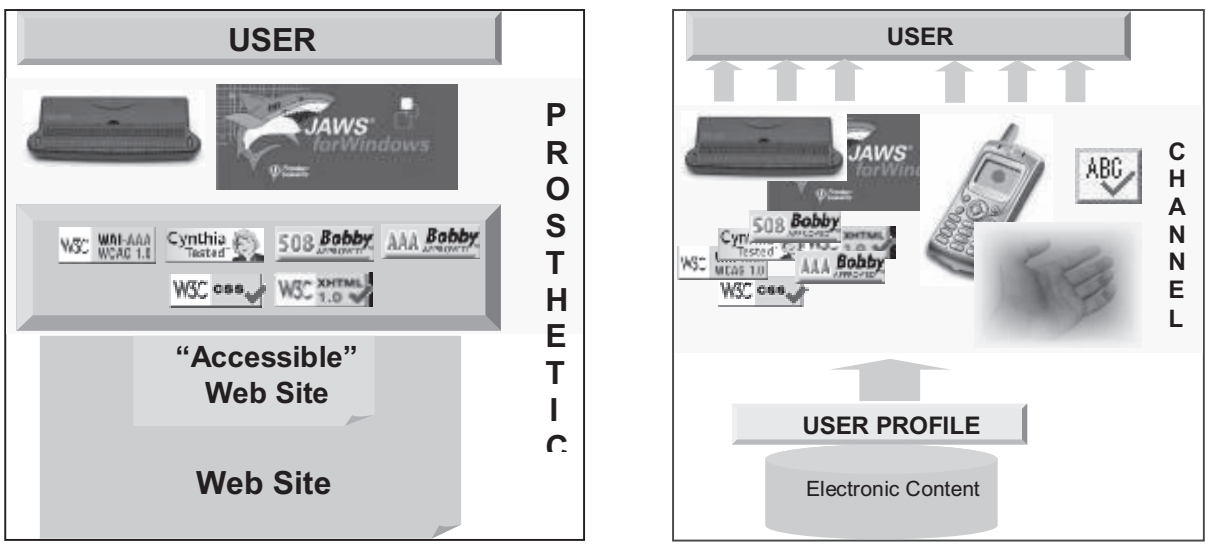

Fig. 1 – *Focusing on electronic contents accessibility (right) over accessibility of the "publishing container" (left), it is possible to provide the impaired user with an adaptive, multimodal electronic publishing environment that better adapts to his/her needs, turning enabling technologies which are commonly used as prosthetics into real information channels.*

MultiAbile can be considered a "multimodal electronic publishing platform", in the sense that it is designed to present electronic contents to its users using different channels and in different forms, according to the defined user profile. These forms are called "modalities" (and the platform "multimodal"), because content can be in principle accessed using more than one of its forms of representation at once [6]. Currently the channels supported by the implemented prototype are the web and the telephone, with possibility to transcode dynamically (with an XSLT-based approach) electronic contents to AAA-compliant HTML readable with Internet Browser, a Browser Screen Reader, or a Braille Display; VoiceXML with TTS/ASR navigation system accessible over the phone; Macromedia Flash cards with pictures, animations and colorful layouts. Over these channels, the user profile determines a number of *modalities* with which content is aptly optimized to better suit the user's needs [10].

These modalities include:

- different levels of language and text structure complexity for the same content, measured with a cognitive complexity index (refined and tested for the moment only for Italian);
- coupling of information with audio and vibro-tactile cues to support navigation into content and exploration of contents of a graphical nature;
- use of mouse/hand gestures to improve or facilitate navigation into content;
- reading with high-quality synthesized TTS (Text-To-Speech) of single sentences or paragraphs.

These modalities, most of which can be used simultaneously mainly using the web channel, are designed to help people with different impairments finding the way which suits them most, in accessing the same structured electronic content. In our early tests, we are validating the benefits of a multimodal publishing environment with:

- visually impaired users, which benefit from using the TTS functionalities, the audio/tactile cues and the Braille display;
- hearing impaired users, which benefit from the availability of a simplified version of the text (in structure and lexicon) which helps to overcome difficulties with the written word;
- dyslexic users, which benefit from the possibility to have the structurally simplified version read by a high quality TTS system;
- users with cognitive disorders (problems in attention or sequencing), which may benefit from a highly graphical, multimedia rendition of the content using Macromedia Flash cards.

2 A Reference Model for the User

Impaired users have been traditionally grouped into profiled categories: blind, deaf, physically impaired, and users with learning disabilities. This approach has had its effectiveness, focusing on specific needs. Nevertheless, this approach has been growingly criticized by associations of disabled people for its narrow vision and for the approach with sets the impaired person in a markedly separated context. During these years have been worked out different ways to represent the situation of disability. In the 1980s WHO's ICDH2 proposed a linear approach distinguishing among impairment, disability and handicap, but still conserving a causal approach. Confrontations between social, medical and associative institutions have produced the ICF, [5] International Classification of Functioning disability and health. This document has the great merit to introduce a more integrated representation of the health state of a person; its four areas are: body functions and structure, activity and participation and environmental factors. ICF has been thought as an operative tool, useful for comparing different levels of intervention both on individual and social politics level and for evaluating the dynamic of the persons' functions. We have taken inspiration from this approach in designing Multiabile's user profile, using attributes borrowed from ICF for representing functional attributes of the user.
This choice has the advantage of focusing on a relatively neutral description of the characteristics of the user, a description that may be adapted to any user, and not specifically to impaired people. This is perceived ,and indeed it is, as an important factor of social integration. The problematic side of this approach is about the availability of all information needed for a complete profiling, partially related to privacy aspects. Obviously the eagerness of the user to supply information about him/herself is related to the perceived benefits from accessing and using the information publishing system. To increase this perception, users may express a set of preferences, both about device availability and about his/her intentions to use, for example, a multimodal or mono-modal interface, the guided or fluent dialogue version of vocal interface.

Fig. 2 – An example on how the user can set his/her profile in the publishing environment to configure multimodal access to electronic content

3 Content Structure and Creation Workflow

To create and organize content in a way that could be dynamically adapted to the users' profile, we adopted a predefined content structure and an electronic content management workflow based on XML and XSLT, with a final step consisting in packaging learning contents into modules according to the Content Aggregation Model provided for in the SCORM standard [1]. The general purpose of the content development framework originally was to allow rapid development of e-learning content modules. Describing electronic content in an XML format allows a complete separation between the content itself and its presentation, making it possible to present the same electronic content in different contexts or on different media. During the production of the first adaptable content, we designed a streamlined chain of activities to obtain a controlled production flow, with a high production throughput.

The general workflow can be described as follows: usually the original contents are supplied by authors in different forms (usually a plain Word document). Starting from this plain document, an editorial curator creates a "storyboard template", reorganizing the provided content marking titles, reorganizing in paragraphs, deciding the general layout of the content for presentation in multimedia format, classifying sentences as more relevant and less relevant (for content restructuring), proposing synonyms and simplified versions of sentences, etc.
This "storyboard" is designed keeping as reference specific "semantic placeholders" that are related to the elements present in the XML data structure, defined in a Document Type Definition. The DTD is organized in a way to be kept as simple as possible (to maintain an error-free and effective content management), but still to be able to capture all necessary content structure information to allow a correct repurposing of the content across channels and modalities. In a second step, content editors receive the storyboard (usually a text document) and proceed creating an XML version of the content. The creation consists in starting from a new XML template using a normal XML editor (e.g. Altova's XML Spy), then the storyboard placeholders are mapped into XML elements. Starting from the storyboard, the XML content module is generated in at least two versions, one using original texts and contents, and another simplified in language, lexicon and structure. The language simplification is verified against a language complexity index, which we have elaborated combining different complexity analysis techniques (the GULPEase index, combined with the use of words taken from a Base Vocabulary of the language and a complexity analysis at morpho-syntactical level) already existing for the Italian language [9].

The following DTD shows the data elements which are used to organize the content.

```
<?xml version="1.0" encoding="UTF-8"?>
<!--DTD generated by XMLSPY v2004 rel. 3 U
(http://www.xmlspy.com)-->
<!ELEMENT game EMPTY>
<!ATTLIST game
        path CDATA #REQUIRED
        visible CDATA #REQUIRED
>
<!ELEMENT item EMPTY>
<!ATTLIST item
       id (voce1 | voce2 | voce3 |
voce4 | voce5 | voce6) #REQUIRED
       status CDATA #REQUIRED
>
<!ELEMENT module (title, subtitle)>
<!ELEMENT page_content (#PCDATA)
(emphasize)>
<!ELEMENT page_title (#PCDATA)>
<!ELEMENT quiz EMPTY>
<!ATTLIST quiz
       path CDATA #REQUIRED
       visible CDATA #REQUIRED
>
<!ELEMENT sco (module, sco_name,
sco_content, sco_menu)>
<!ATTLIST sco
       version CDATA #REQUIRED
>
<!ELEMENT sco_content (sco_page+)>
<!ELEMENT sco_menu (item+)>
<!ELEMENT sco_name (title)>
<!ELEMENT sco_page (page_title,
page_content, quiz, game)>
<!ATTLIST sco_page
       id (1 | 2 | 3 | 4 | 5 | 6)
#REQUIRED
>
<!ELEMENT subtitle (#PCDATA)>
<!ELEMENT title (#PCDATA)>
```

When the document is complete, it is validated against the DTD: this ensures that the document is both well formed and valid. In this developing process, it is high unlikely that errors can occur without generating some validation errors. The XML editor assists the developer in creating structured content and the conformance between the presentation template adopted by each channel and DTD structure description enables an easy flow from content to presentation, maintaining a clear separation that makes possible the reuse of the XML content files in other applications/publishing environments. Once content is structured following the present DTD, it can be transformed using an XSLT transformation stylesheet, using a predefined template into which the content itself will presented (e.g. a graphical frame for Flash presentations, a template web page for HTML, a voice navigation prompt system for VoiceXML, a standard page layout for PDF documents, etc.).

Once content is transformed into one or more version, the third step in the process is to package it as a learning object (commonly called SCO in the SCORM course structure) into a single deliverable SCORM package. This SCORM package is ready to be installed into a regular e-Learning platform, or into a publishing environment which allows multimodal interaction with the content itself, such as the MultiAbile site.

4 Content packaging and transformation

The aggregation of the learning content is made by creating an XML document named “imsmanifest.xml” in which one defines the organization of the module in sections, lessons, quizzes and so on, and specify all resource files needed to use the course (XML files with the actual content of each learning object, images and so on). All learning objects and course assets are finally archived together with imsmanifest.xml in a compressed ZIP file that can be uploaded to the e-learning platform that manages the contents. The course is thus published for registered users - keeping track of user access, session time, lesson location, quiz scores and so on, according to the SCORM v1.2 “Runtime Environment Specification”. Using XML as the basic language to define content is the most convenient way to deliver the same course among different channels and to provide to the user different modalities of interaction with the system. One of the most interesting modality of interaction provided by the MultiAbile LMS is the vocal user interface realized using VoiceXML, that is delivered through a voice gateway that gives access to the platform by using a normal telephone. Content is synthesized by a special software and user can interact with the system using his/her own voice and DTMF codes by typing on the telephone keypad. The vocal interface is created by transforming learning objects contained in XML files and imsmanifest.xml using XSLT. The VoiceXML document created in such a way provides several commands to make the interface more accessible: a user can listen to the content, going back and forward through paragraphs, repeat the current paragraph, pause and restart reading, and so on. The manifest file provides the user with an index of the course and an easy way to access learning objects. The LMS tracks user activity so that he/she could switch between graphical and vocal interface keeping his/her completion status and scores.