Leveraging Rich Accessible Documents on the Web

Rui Lopes
LaSIGE/University of Lisbon
Campo Grande, Edifício C6
1749-016 Lisboa, Portugal
rlopes@di.fc.ul.pt

Luís Carriço
LaSIGE/University of Lisbon
Campo Grande, Edifício C6
1749-016 Lisboa, Portugal
lmc@di.fc.ul.pt

ABSTRACT

This paper presents a new approach on leveraging accessibility for rich document delivery to the Web. The proposal entails a profile modeling task, where multidisciplinary teams can discuss users, devices, and usage scenarios, in order to grasp and synthesise the different document delivery scenarios. A document production framework is presented, which can be configured according to modeled profiles. By using this approach, documents are tailored to users in such a way that rich interaction capabilities are maintained, without sacrificing content accessibility.

Categories and Subject Descriptors

H.5.4 [**Information Interfaces and Presentation**]: Hypertext/Hypermedia—*User issues*; K.4.2 [**Computers and Society**]: Social Issues—*Assistive technologies for persons with disabilities*; I.7.2 [**Document and Text Processing**]: Document Preparation—*Hypertext/hypermedia, Multi/mixed media*

General Terms

Design, Human Factors.

Keywords

Web Accessibility, Document Production, Profile Modeling

1. INTRODUCTION

Accessibility on the Web is being taken more seriously as time goes by. Governments have started to legislate towards info-inclusion, resulting on an increased awareness of accessibility. Several guidelines provide helpful cues for creating contents that have no barriers on accessibility [7]. However, awareness and guidelines are just entry points on delivering rich acessible documents on the Web. Nowadays, accessibility is still regarded as providing alternate (and overly simplified) versions of contents. This results in limited user experiences from those who rely on accessible contents, in comparison to richer experiences of the average user.

Different realistic scenarios can be regarded as specialisations of accessibility. Physical disabilities, such as the blind or the deaf, require that document formats do not compromise information access. Cognitive disabilities, as present in the elderly, require a special attention on navigation and interaction aspects within documents. On the other hand, the Mobile Web initiative[1] thrives for accessing information everywhere, therefore devices capabilities and limitations may pose significant accessibility problems to users. Furthermore, when putting the user and devices in specific usage situations (e.g., outdoors), additional problems may arise when accessing documents within a Web browser. Summarising, accessibility must be perceived nowadays as a crosscutting concern over users, devices, and usage situations.

At the same time, users are starting to demand for richer interaction capabilities on the Web. Users want to interact with documents by annotating them, sharing opinions with others, viewing them in mobile devices, etc. Browsers are starting to provide the required technological support, such as being available in different devices and affording richer scripted behaviours. Hence, they can be exploited on document production scenarios, in order to deliver rich documents in any device and usage situation, without compromising users' accessibility requirements.

However, producing documents to the Web requires a multidisciplinary approach. Developers, graphical artists, managers, usability experts, and others, need to grasp these rich scenarios in such a way that accessibility is not compromised, and user experiences are improved. Therefore, a high level approach on understanding users, devices, and usage situations should be available to teams and, at the same time, tightly coped with document production scenarios. In order to support them, this paper presents a vision of rich accessible documents on the Web based on the definition of profiles for users, devices, and usage situations with a modelation methodology incorporating user modeling ontologies. A supportive document production framework is also presented, affording different tailoring tasks over rich accessible documents.

2. RICH ACCESSIBLE DOCUMENTS

As of today, the information found on the Web can be regarded either as structured or semi-structured. While

W4A2007 - Technical Paper, May 07-08, 2007, Banff, Canada. Co-Located with the 16th International World Wide Web Conference.

[1] http://www.w3.org/Mobile/

structured data is perceived as stable tabular information, semi-structured data does not fit this scenario. When data order is crucial, and structure is volatile and relatively unpredictable, we are faced with document-centred scenarios. This intrinsic characteristic poses several challenges on providing an accessible, yet rich and interactive, way of delivering documents to users.

Users interact with documents through reading activities, which are greatly influenced by their goals. Reading a novel for entertainment purposes, or studying a textbook, are reading activities that engage the reader with different levels of commitment and attention. To portray the whole spectrum of reading situations, a categorisation was proposed in [30]. Two dimensions – nature of engagement and breadth of the activity – are used to categorise four conventional reading situations:

- *Passively reading a text*, often associated with entertainment reading, occurs, for instance, when reading a novel;
- *Passively reading multiple texts*, occurs when reading e-mail messages. This category is associated with keeping informed;
- *Actively reading a single text*, is associated with learning tasks, as in studying a textbook;
- *Actively reading multiple texts*, is often associated with research activities, where information is gathered from multiple sources.

While understanding the text is a common goal for all reading situations, different situations will pose different problems. Situations encompassing multiple texts entail the need to manage multiple documents and the difficulty of finding the needed information. Active reading [1] involves underlining, highlighting and annotating, either on the text or separately, thus demanding for some mechanism to manage annotations.

Digital documents and Websites contribute to mitigate some of these problems. Digital documents offer the possibility to record, organise and search annotations entered by the reader [29]. Websites leverage the management of large document collections and prosper the creation and exploration of relations between documents. Moreover, the possibility of sharing personal annotations within the community of members of a Website (such as a digital library) conveys a social sense to documents [18].

Besides these affordances, the document's digital support opens up the possibility of enriching its content with supporting media [8]. If allowed by browsers, the document can have its content enriched with additional multimedia content, like images, videos, music, and sounds. In such a platform, the document's content can also be narrated in addition or in alternative to the visual presentation, similarly to digital talking books [11]. Coped with active reading perspectives, a document will start to be perceived as a hybrid between its intrinsic content and structure, and a reading platform.

Furthermore, the rich accessible document in a Web context is an important step to reach a heterogeneous audience and a variety of situations, as follows:

- Researchers, students and other professionals who can benefit from advanced navigation features, annotation support, and the integration with a large collection of related material;
- Pleasure readers in constrained environments can benefit from alternative presentation modes. An audio-centered document can be used in situations where visual focusing is cumbersome (e.g., on a mobile phone);
- Both children and adults can enjoy visual and audio enhancements that enrich and complement the document's presentation;
- Children and other reading and writing disabled persons can benefit from simultaneous narration and visual presentation when learning how to read and write;
- Visually impaired users can benefit from audio narration, coupled with navigation mechanisms (e.g., table of contents), and speech-based annotations;
- Partially sighted users will benefit from customised visual components (e.g., increased font sizes and tailored color schemes);
- Cognitive disabled users require adapted contents and interaction capabilities in order to achieve several goals (e.g., simplified content, reduced navigation tasks).

Inherent to these issues come the questions about output and input modalities, and how they fit into an accessible Web scenario. Regarding rich documents' output capabilities (i.e., presentation) they can be perceived as critical (e.g., appropriate document formats for the blind), or as improvements over traditional characteristics (e.g., simultaneous narration). Consequently, the classification of the act of reading can be augmented according to users' perceptions, in order to better adapt to their accessibility requirements, as follows:

- *Textual:* this is the traditional form of reading, as the document is presented in a textual format, suitable to several users and scenarios;
- *Graphical:* when the cognitive effort required to grasp a given document is high (e.g., incapability of reading text), a graphical counterpart of the document may leverage its comprehension. Consequently, this reading situation will be centred around the visual interpretation of graphics (e.g., images, videos);
- *Aural:* visual stressing situations will require an audio counterpart of a document (whether in form of a speech track or real-time synthesised speech), in order to lower cognitive stress. Thus, reading will be perceived in an audio listening scenario.

In what concerns interacting with a rich accessible document, input modalities must be also selected appropriately, in order to cope with users' accessibility requirements and devices' constraints. Moreover, special care must be taken when documents are tailored into active reading scenarios, i.e., to provide input modalities that will leverage tasks such as annotating or navigating the document through different paths without compromising usability.

When multiple input and output modalities are combined, documents will have richer capabilities allowing for richer scenarios to be available to users (aural output with a voice input, for instance). However, this empowerment of rich accessible documents should be used with caution, in order to keep users' focuses into reading activities, not the opposite. Otherwise, an overwhelming cognitive effort from users will dismiss the benefits of rich accessible documents.

While all these issues will be ultimately reflected on implementation details, the Web's nature requires a multidisciplinary approach to be taken into account at early stages of designing document centred delivery scenarios (such as Websites). Therefore, a high level discussion centred on designing rich and interactive, yet accessible, documents for the Web should be comprehensible for developers, graphical artists, usability experts, accessibility engineers, and project managers.

As the authoring process for accessible documents is also perceived as a multidisciplinary task, manual processes are typically used in order to cope with accessibility and usability (e.g., linking a document's structure with its speech based audio counterpart). Too much effort is required for scaling up the availability of rich accessible documents on the Web and, most importantly, errors become easy to introduce within and between documents. Adding up the complexity and richness of the whole spectrum of scenarios described previously, delivering rich accessible documents becomes an impossible goal to achieve. Therefore, an automated approach should mitigate such problems.

The large spectrum of possibilities presented will require a framework where the different intervenients on document authoring processes can discuss users and usage situations at a high level and, at the same time, take advantage of this knowledge on creating and configuring automated authoring environments. This way, rich acessible documents will be always tailored to users in a coherent way. Such modeling task is presented on the next section.

3. PROFILE MODELING

For the purpose of getting a better perception of user, device and usage situation characterisations, the concept of profile modeling emerges. In a nutshell, a profile reflects a document interaction scenario, by grouping miscellaneous characterisations in a systematic way. In order to cope with the broad range of scenarios previously discussed, a profile modelation task should meet the following requirements:

- *High-level viewability:* profile modeling should be easily grasped by experts in different areas. It should be easy for non-developers experts to model profiles without having to learn low-level constructs (e.g., scripting, expert shells, etc.). Moreover, a graphical presentation of profiles will better adequate to plan and discuss different document delivery scenarios on blackboards, thus promoting interactions within teams;

- *Foster reuse:* a profile metamodel should supply a set of constructs that leverage their reuse across different scenarios. With these mechanisms, the similarities and differences between scenarios can be exploited in such a way that coherence is leveraged (i.e., sharing concepts) and particularities emphasised (i.e., tailoring);

- *Direct applicability:* if profile modeling is excessively informal, it will only have the purpose of guiding developers on document production implementation tasks. Consequently, incoherences between models and their reflection on production frameworks may yeld critical accessibility problems on document delivery, decreasing the acceptance ratio from end users. Therefore, profile modeling should have a direct applicability on production frameworks, e.g., by using formal concepts from ontologies. Moreover, with this formalisation, profiles can be integrated in adaptive Web platforms.

Based on these assumptions, a high level modelation tool has been created with the Eclipse Modeling Framework [6]. Within this environment, a profile metamodel was specified, as depicted on Figure 1:

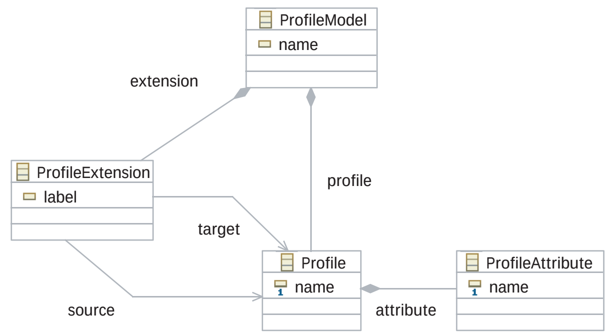

Figure 1: Profile modeling metamodel

At its root, a profile model (*ProfileModel*) is composed by a set of classes (*Profile*) and their associated attributes (*ProfileAttribute*). Each class will describe a particular profile, according to the scenario-specific requirements gathered earlier, e.g., through contextual design methodologies [4]. Each attribute will be a representation of a concept that characterises its enclosing profile (e.g., an ontology class). It is worth mentioning that specifying attributes for profiles is based on an open world assumption, i.e., what is not stated on a profile is not necessary false within the profile's domain, it is simply unknown.

With these basic constructs any number of profiles can be defined towards miscellaneous document processing domains. However, as profile count increases, the probability of having profiles sharing concepts rises. Consequently, the profile metamodel introduces a specialisation mechanism (*ProfileExtension*) to specify inheritance-based relations between profiles. This way, profiles that share a set of attributes can be refactored to inherit these from a parent profile. Moreover, profiles afford multiple specialisations (i.e., inheriting attributes from more than one profile), allowing the modelation of more complex scenarios.

Going deeper on profile characterisation (i.e., attribute definition), two approaches can be envisioned: informal *vs.* formal. The former allows for a coarse description of concepts (e.g., a short text or list of keywords) and can be understood at a high level by the different intervenients on the modeling process. However, having a direct applicability is critically required, as previously explained. Therefore, informal characterisations should be either translated into formal definitions, or promptly formally specified from the beginning, at the expense of lowering the profiles' expressiveness.

Some off-the-shelf ontologies provide a significant groundwork for characterising profiles for users, devices, and usage scenarios. GUMO [16] provides a thorough research on user modeling in the form of an OWL [31] ontology, with a set of concepts ranging from abilities (e.g., *AbilityToHear*) to emotional states (e.g., *Satisfied*). On device characterisation, CC/PP [37] (Composite Capability/Preference Profiles) provides an RDF [3] framework to define different device profiles (named *components*). However, being just a framework, it does not provide the required richness of device charaterisation instances to be promptly used in profile modeling tasks. UbisWorld [15], an ontology that complements GUMO's concepts, provides high-level concepts for ubiquity scenarios such as devices (*Handheld*, *Mouse*, etc.), vehicles (*Bus*, *Airplane*, etc.), spatial purposes and locations (*Leisure*, *AirportParis*, etc.). As stated previously, these ontologies just provide bootstrap concepts to be used on profile modeling tasks. Nevertheless, ontology engineering practices [32] can leverage particular concepts that may make sense to be used in specific profile modeling instances.

To illustrate the modelation of different profiles within a rich accessible document domain, the following scenario is presented: a public library must make available different rich accessible documents to casual readers, whether visually impaired or nonimpaired. Moreover, to stimulate the usage of the digital medium, an ubiquitous access to the documents should be supported (at least for nonimpaired audiences). Based on these requirements, the profiles *Casual Reader*, *Blind*, *Desktop*, and *Mobile* were defined, using the GUMO and UbisWorld ontologies, as presented on Figure 2:

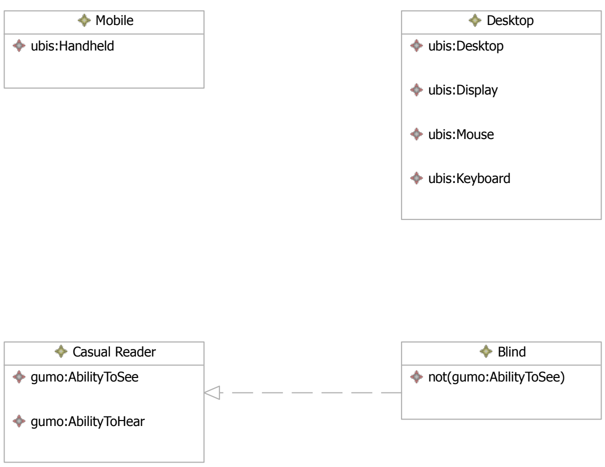

Figure 2: Profile model for a public library scenario

This model defines the *Desktop* profile as a conjunction of ontology concepts that characterise a typical desktop computer environment (i.e., display, mouse, keyboard), as opposed to the scarce knowledge available for general mobile environments as presented on the *Mobile* profile. Regarding users, the *Casual Reader* profile defines this user class as people who has the ability to see and hear. The *Blind* profile emerges as a derivation of *Casual Reader* (depicted as an arrow pointing to the parent profile), but stating an inability to see (*not(gumo:AbilityToSee)*). However, as these profiles were modeled in a disconnected fashion, the particularities of usage situations (e.g., a blind person using a desktop) are not taken into account. Consequently, a derivation from this model is depicted on Figure 3, encompassing the usage situations described in the scenario.

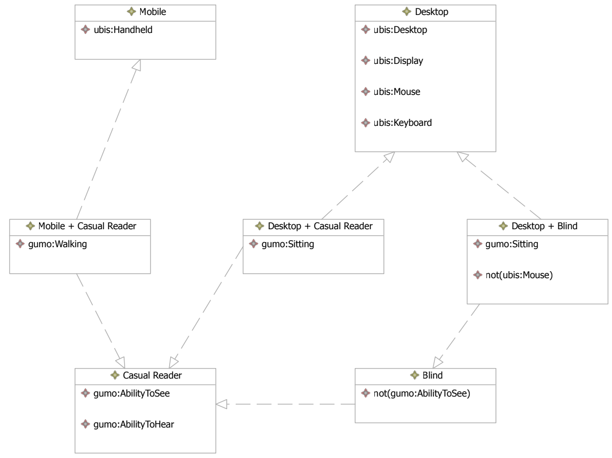

Figure 3: Profile model with usage situations

Here, three usage situations provide a better undestanding of the limitations they impose. The *Desktop + Casual Reader* profile states that the user is sitting (i.e., reduced mobility). This type of knowledge can be used at latter stages in document production frameworks, for instance, to add richer contents that require a higher attention level from users (e.g., a video). Regarding the *Mobile + Casual Reader* profile, it contrasts with the previous profile from the attribute *gumo:Walking*, that may imply high mobility from users. Once more, document production frameworks may use this information to transform documents towards *aural reading* scenarios. Lastly, the *Desktop + Blind* profile states that mouse-aware interaction should not be taken for granted, as this device imposes severe accessibility issues to the blind.

It should be stated that, despite the simplicity of this scenario, profile modeling may become complex in order to cope with richer scenarios. This yelds for a set of good practices that should be taken into account when engineering profiles for rich accessible documents on the Web. However, it is out of scope of this paper to define such good practices. Next, a production framework for rich accessible documents is presented.

4. PRODUCTION FRAMEWORK

To cope with such a rich user audience, range of devices, and usage situations that profile modeling tasks are able to grasp, a flexible production framework has been created, based on early results from automated rich document processing [24, 25]. This flexibility is reflected both in the document format used at its base, and in its processing components. The key aspect of this framework concerns on creating reusable document processing components that can be recombined. This allows the production of documents according to the previously modeled profiles. The inheritance mechanism of profile modeling affords the necessary cues to grasp, later on, which components may be reused accross different production framework configurations.

This framework was implemented with XML pipeline architectures [5, 22], bridging processing components implemented with XSLT [19] and special purpose Java classes. Each framework configuration is specified as an application of pipeline technologies, by selecting appropriate processing components according to modeled profiles. This way, project managers can coordinate their teams according to the processing components required to support the profiles (e.g., teaming a usability expert with a developer to create a navigation scheme optimised for *Mobile* based profiles).

Going deeper on the production framework, a separation of concerns groups processing components according to their processing semantics, as depicted on Figure 4.

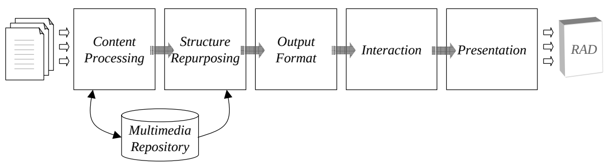

Figure 4: Production framework overview

The framework receives on its input an initial content (e.g., any document), which is processed by several steps along a specific configuration of the different concerns (in compliance with the modeled profiles), resulting on a rich accessible document ready to be consumed by users. A multimedia repository provides additional support to the framework, allowing the inclusion of external contents in documents (e.g., additional pictures to illustrate some topic), and storing the documents for content reuse scenarios.

From the previous sections, it is fair to make the assumption that rich accessible documents can be composed by static media, dynamic media (i.e., time-based content such as audio and video), or even a mix of both. Despite the base medium of rich accessible documents, their structures (e.g., chapter, section, paragraph, etc.) are highly stable. Therefore, DocBook [38] has been embraced by the production framework as the basis for describing documents and processing them with the framework's components.

However, as DocBook by itself does not provide constructs to cope with time-based content (such as an audio track composed by the speech counterpart of a document), the format has been extended to support such scenarios. As an example, Figure 5 presents an excerpt of the structural definition of an audio track.

```
<media>
  <item id="a1" begin="0"
        end="0.3" expr="The" />
  <item id="a2" begin="0.35"
        end="0.77" expr="quick" />
  ...
</media>
```

Figure 5: Audio structure definition

This description allows the identification of speech excerpts within the audio track timeline (with the `begin` and `end` attributes), as well their textual representations (`expr`). These expressions cope with the possible differences between a document's text and its audio counterpart (e.g., identifying *2007* in the text vs. *two thousand and seven* in the audio). Appropriate descriptions were conceived also for video and images, in order to identify documents excerpts within different media files.

However, having unrelated representations of a same content is not sufficient, specially as non textual document sources do not have semantic information about the document's structure (therefore decreasing significantly the interaction capabilities of a rich acessible document). Consequently, the production framework affords this type of relations based on XLink [12], according to the exemplificative syntax presented on Figure 6. More complex content relationship scenarios are also supported, based on XLink's extended link capabilities.

```
<relations>
  <arc xlink:from="w1" xlink:to="a1" />
  <arc xlink:from="w2" xlink:to="a2" />
  ...
</relations>
```

Figure 6: Relations definition

Despite the inherent power of using this type of document structures, extracting this type of information is not trivial. In the case of audio, translating speech to text must be done with high accuracy levels. If this is not accomplished, the discrepancies between audio and textual content of a document will repercute later on the resulting document, yelding significative accessibility problems (e.g., cognitive pressure induced by synchronisation failure between audio and text presentation).

Next, the implemented processing components are described for each framework concern. It is worth mentioning that, being just a prototype, only a handful of components is available. Nevertheless, other components can be easily integrated in the framework seamlessly, in order to support different rich accessible interaction scenarios.

4.1 Content Processing

The first concern within the document processing framework relates to tasks centred on content processing, such as content normalisation and reasoning, which can be used in any framework configuration.

Scaling up this type of document processing frameworks can not be done without coping with already existing corpora. However, documents are stored in a wide range of formats (such as PDF, Microsoft Word, audio tracks, etc.) and must be normalised into DocBook. To accomplish this task automatically, the production framework has a preprocessing step based on the XPR language (XML Pipeline Rules) [23], where document format sensors and transformation rules afford the required normalisation.

After this task, multimedia content reasoning is applied to the normalised documents and their multimedia counterparts, according to the multimedia content classification architecture proposed in [25]. In a nutshell, reasoning is performed with feature extraction and classification algorithms, to be indexed and stored in the multimedia repository. As an example, when speech-based audio tracks are fed to the framework, this task applies speech alignment algorithms to extract the relations between audio and textual counterparts of documents, using the mechanisms presented in [13]. This way, the profiles centred around *aural reading* situations (such as *Mobile + Casual Reader* and *Desktop + Blind*) will have richer content navigation capabilities.

4.2 Structure Repurposing

At this concern different repurposing tasks can be applied, such as extracting navigation paths, document simplification, and perform chunking/pagination on documents.

Regarding the extraction of navigation paths, the production framework affords this task in order to provide users with alternatives on navigating through documents, such as extracting the table of contents, sidenote lists, or even indexes. This is performed by looking at the document's DocBook structure and extrapolate it into separate documents (e.g., by seeking `chapter`, `section`, and similar DocBook elements) or by applying "back of the book index" building algorithms [9]. Having all these alternate navigation paths opens the way for richer experiences from users on interacting with documents, without sacrificing accessibility concerns.

Another processing task relates to document structures' simplification. Documents targeted to the Web are traditionally created to be read on desktop environments, therefore assuming a significant hardware perfomance. However, when talking about computational limited devices (such as Web enabled mobile phones), these perfomance requirements must not be overlooked. One of the ways this can be tuned is by reducing the complexity of document structures (e.g., removing `phrase` element tags from DocBook structures) and associated XLink relations. Also, essentiality tracks techniques [2] can be further applied to adapt content to user's requirements.

Lastly, this concern can also be used to perform chunking and pagination tasks over documents. This will allow for an improved adequation to resource-limited usage scenarios such as *Mobile* based profiles, or to convey a more physical sense of documents (as traditional Web documents can fit into just one page).

4.3 Output Format

At some point in the document processing flow, documents have to be translated into formats that are understandable by Web browsers and related applications. This aspect is handled by the *Output Format* concern within the production framework.

Choosing an appropriate output format must take into account the different factors inherent from the profiles modeled earlier. The combination of user, device, and usage scenario dictates which format transformation should be applied to documents. In the scenario presented earlier, for instance, *Desktop* based profiles can benefit from the transformation of documents into the XHTML+SMIL format [26], where plain HTML contents can be integrated with advanced multimedia syncronisation capabilities. This way, situations such as *aural reading* are afforded on traditional browsing experiences. Other options relate to the transformation of document contents into SMIL [10] presentations, in order to cope with *aural reading* mobile scenarios, such as the *Mobile + Casual Reader* profile. A set of processing components were created, in order to transform docuemtns into the described formats.

In order to support active reading, the documents must be augmented with adequate mechanisms (such as attached scripts, links to trigger activities, etc.). Therefore, the output format concern provides another processing task for this purpose. However, it should be noticed that document output formats may have inherent limitations (e.g., such as SMIL's lack of scripting capabilities). Therefore, chosen output formats may have to be reconsidered if active reading scenarios have to be supported.

4.4 Interaction and Presentation

After transforming the document into an output format targeted at a specific profile, interaction and presentation capabilities must be introduced, in order to deliver the document with high accessibility compliances and overall quality. Therefore, the interaction and presentation concerns are introduced in the document production framework.

Regarding interaction, different tasks can embed input device specific capabilities in the document. Such capabilities can be, for instance, producing a vocabulary for speech recognition components, limited to the set of words found in the document. This way, speech recognition accuracy may be increased and used for speech based interaction with documents (e.g., searching). Other interaction features that enrich documents without sacrificing accessibility and usability, relate to the cases where mouse interaction (if a profile model allows it) leverages navigation tasks. With this type of interaction, point and click interaction can be exploited, (e.g., click on a page preview jumps directly to the appropriate page), in comparison with the discrete interaction afforded by keyboards.

Once again, this type of capabilities are only supported when the selected document output format allows their specification. For instance, as Internet Explorer allows embedding ActiveX components, speech recognition can be introduced on documents without requiring native support of this input modality on the browser, whereas most SMIL players do not allow the adaptation of interaction mechanisms, leaving those capabilities to their enclosing runtime environments. Such considerations are taken into account by the processing tasks available at the interaction concern.

The last concern introduces a set of tasks that adequate the presentation of documents to users. This can include applying appropriate stylesheets to users (e.g., children require different document presentation from the eldery), or complying with specific branding requirements (e.g., logotypes, font selection), or selecting aural icons to identify document structures on *aural reading* situations. In order to provide a coherent look and feel to documents across modeled profiles, all these aspects are configured at a high level and subsequently translated into each profile browsing platform accordingly. This way, for instance, graphical artists can focus their work on providing the necessary assets for documents without being tied to particular output formats.

5. RESULTS

The profile models presented earlier on this paper served as the basis for the definition of different configurations for the production platform. The three situations (*Desktop + Casual Reader*, *Destkop + Blind*, and *Mobile + Casual Reader*) encompassed different requirements, resulting on the implementation of several processing components. Several of these components were reused across profiles (e.g., normalising content, extracting table of contents), therefore leveraging the effort required to provide tailored documents to different users. But, as expected, some components had to be developed towards just one profile (e.g., adding mouse interaction to XHTML+SMIL capable browsers). Nevertheless, all these components could be used on the cases where usage situations were not known *a priori* (e.g., *Casual Reader* and *Desktop* profiles), enforcing coherence between all scenarios.

Regarding the *Desktop + Causal Reader* profile, the selected processing components were able to produce a rich accessible document based on an initial input (the *O Senhor Ventura* novel, written by the portuguese writer *Miguel Torga*), both in textual and audio formats. As depicted on Figure 7, the enriched document affords different navigation paths (table of contents and sidenotes) over fully syncronised visual and aural representations of the content, presented on Internet Explorer.

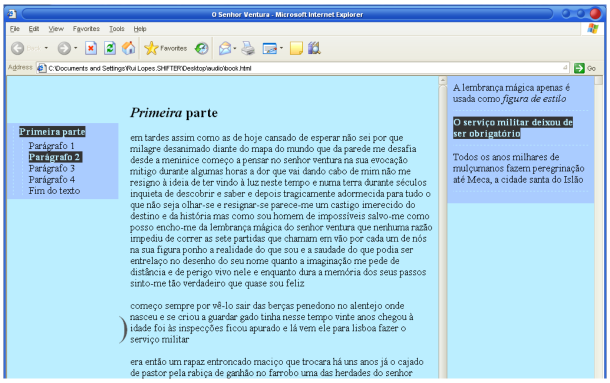

Figure 7: A rich book on Internet Explorer

For the *Mobile + Casual Reader* profile, the same document (*O Senhor Ventura*) was fed to a tailored configuration of the production platform. In this case, structure repurposing components had to be applied in the document transformation pipeline (e.g., reducing document structures, chunking). As the selected usage scenario concerns a *ubis:Walking* concept, an *aural reading* tailoring was applied to the rich document, under the form of a SMIL document output format (as playback is supported in several mobile platforms). This lead to selecting audio as the primary source of content delivery, coupled with a simple visual presentation mechanism, as presented on Figure 8.

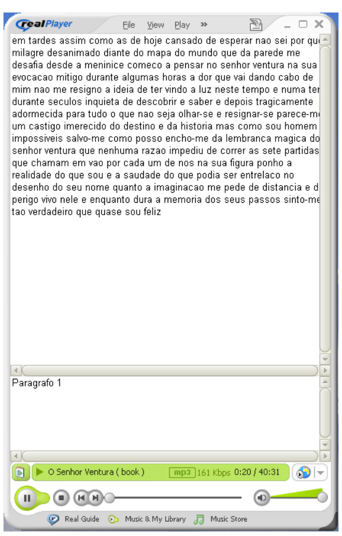

Figure 8: A rich book on Real Player

Lastly, the *Desktop + Blind* profile resulted on selecting a set of components that lowered the cognitive effort required to read the *O Senhor Ventura* novel in an *aural reading* scenario. For instance, by repurposing the document's structure to encompass just *chapters*, aural icons interfere less in the document presentation. As the desktop environment was selected on the profile, the document was produced to be presented also on Internet Explorer (due to its XHTML+SMIL capabilities), but without featuring visual presentation (in comparison to the first profile). Moreover, interaction capabilities were disabled in the document, opening the way to users preferred input devices (as this issue is critical in order to reach a satisfaction level from users).

6. DISCUSSION

As of today, accessibility on the Web is regarded (*when* is regarded) simply as following *ad-hoc* good practices or, at most, WCAG specifications [7]. This typically results on providing simpler versions of documents or Websites, therefore resulting on lower experiences from accessibility dependent users. Furthermore, accessibility issues must not be taken into account just for the disabled. A large range of scenarios require similar approaches, such as device constraints and usage scenarios – the ubiquitous Web. The profile modeling methodology and the production framework presented in this paper contribute to mitigate these issues, by changing the focus of tailoring documents from manual tasks (cumbersome and error-prone) to automated approaches based on high level, yet formalised, concepts. Moreover, as rich accessible documents become available, a higher level of interaction and commitment from users is fostered.

However, the Web *status quo* poses severe limits on providing rich accessible documents everywhere, to everyone. When shifting away from traditional users and desktop environments, a lot of improvements still have to happen. Mobile browsers have limited interaction capabilities (e.g., scripting behaviours that can not be synthesised with declarative markup), and scarce multimedia support (no navigation on audio). Fortunately, initiatives such as Mobile Web or SMIL 3 will allow targeting production frameworks to richer and accessible environments, without compromising user experience.

Going deeper on profile modeling, a lot of work should be done. GUMO, CC/PP, and UbisWorld provide initial concepts upon which profiles can be defined. However, these ontologies are not able to grasp the broad range of users, devices and usage situations that should be taken into account when producing rich accessible documents. For instance, it is impossible to specify profiles based on cognitive disabilities (e.g., colour blind), or device clusters (e.g., mobile phones with XHTML capabilities). An ontology that would be capable of covering these concepts could be used to model richer profiles and, consequently, be used to better support document production frameworks.

7. RELATED WORK

The User Modeling research field[2] has provided a significant effort on understanding users, and how supportive systems can exploit this information to adapt contents to user's requirements and preferences. Such systems, like AVANTI [14], are able to adapt content and navigation towards previously defined user models. However, these systems are typically oriented to adapt Websites (i.e., data ori-

[2]http://www.um.org/

ented applications), therefore are unable to cope with semi-structured documents. Moreover, these systems rely on expert shells to model users, as thoroughly dissecated in [20]. This requires high skills both on creating user models, and even on understanding them, contrasting with the high level profile modelation methodology presented in this paper.

A similar approach to profile modeling and document delivery to the Web is described in [33], where profiles are specified as a list of user related requirements that should be taken into account when producing documents. However, this approach just takes into account user accessibility problems, not device related, neither usage scenarios. This decreases the acceptance of documents in ubiquitous environments. Moreover, the described framework does not cope with rich documents scenarios (e.g., providing alternate or complementary document corpora, richer navigation capabilities, or tailored interactivity).

On producing documents and books specifically tailored to the blind (a highly accessibility dependent audience), several results have been taken into account when defining the production framework presented in this paper. These technologies are mainly centred on the Digital Talking Book specifications, therefore inheriting its benefits and problems. State-of-the-art systems, such as eClipseWriter [17], provide facilities to deliver full blown DTBs to the blind based on text-to-speech technologies. However, these systems' results become less satisfactory, as users tend to dislike the robotised voices in lenghtier amounts of time. TTS technologies also do not cope with miscellaneous ambiguous contents [27] (e.g., *IV* read as *i v* instead of *four*), a problem solved with the production platform presented in this paper, through the use speech alignment technologies. Moreover, as none of these systems are based on user modeling capabilities, the ability to produce documents towards the rich usage scenarios presented in this paper is dismissed.

A framework for producing enriched multimedia content towards the Web has been proposed earlier [36], where several transformation steps are applied in succession on multimedia documents, tailored to different usage situations. The system also encompasses a user model centred transformation step [21] to adapt documents to miscellaneous user requirements. However, these mechanisms are used towards personalisation scenarios (not accessibility). Moreover, being centred on multimedia, the system loses its focus on document structures, therefore removing significative interaction capabilities from users (e.g., no table of contents).

The Web Engineering research field also provides helpful cues on delivering highly accessible information on the Web. One particular approach, WSDM [35], defines a user-centred approach for creating Websites. It suports user modeling [34], albeit a fairly limited mechanism. More recently, this methodology has been extended to support accessibility specific requirements [28], where application specifications are automatically annotated for accessibility scenarios. However, this approach is highly centred on data models, thus not fitting document based scenarios. Other limitiations of WSDM, in the context of this paper, relate to its inadequacy towards *aural reading* scenarios, as it is mainly centred in textual information.

8. CONCLUDING REMARKS

This paper presents a novel approach on producing rich accessible documents for the Web. This approach is based on modeling *profiles* as a conjunction of user, device, and usage situations that must be taken into account in order to provide tailored documents to end users. To cope with the huge range of possible scenarios that can be described with profiles, a production framework has been created. This framework was subdivided into different concerns (e.g., structure repurposing) which can be configured according to any modeled profile. As a result, the framework outputs tailored documents, where users are faced with higher levels of interaction, whether they have critical disabilities, or other forms of content accessibility restrictions. Consequently, documents can be transformed into *aural* or *graphical* representations of the same content, enriched with complementary contents, improved navigation capabilities, and even tailored interactivity.

The work presented here has a lot of room for improvement. The profile modeling methodology and the production framework must be further augmented towards richer scenarios as the Web evolves, including:

- Define an ontology that covers accessibility scenarios (such as users, devices, and usage scenarios);
- Adequate the production framework to automatically take advantage of modeled profiles;
- Affording the visual configuration of the production framework, fitting the multidisciplinary nature of creating Web content;
- Establish a set of good practices on profile modeling (e.g., design patterns, a central repository for storing profiles);
- Improve the enrichment of accessible documents, coping with the latest developments of Web technologies (e.g., AJAX, Adobe Apollo).

9. ACKNOWLEDGEMENTS

This work is being funded by Fundação para a Ciência e Tecnologia, through grant POSI/EIA/61042/2004 and scholarship SFRH/BD/29150/2006.

10. REFERENCES

[1] M. J. Adler and C. V. Doren. *How to Read a Book*. Simon and Schuster, New York, 1972.

[2] M. T. Atkinson, J. Dhiensa, and C. H. C. Machin. Opening up Access to Online Documents using Essentiality Tracks. In *W4A: Proceedings of the 2006 international cross-disciplinary workshop on Web accessibility (W4A)*, pages 6–13, New York, NY, USA, 2006. ACM Press.

[3] D. Beckett. RDF/XML Syntax Specification (Revised). `http://www.w3.org/TR/rdf-syntax-grammar/`, February 2004.

[4] H. Beyer and K. Holtzblatt. *Contextual Design: A Customer-Centered Approach to Systems Designs*. Morgan Kaufmann, 1998.

[5] E. Bruchez and A. Vernet. XML Pipeline Language (XPL) Version 1.0 (Draft). W3C Member Submission, World Wide Web Consortium (W3C), 2005. `http://www.w3.org/Submission/xpl`.

[6] F. Budinsky, D. Steinberg, E. Merks, R. Ellersick, and T. Grose. *Eclipse Modeling Framework*. Addison Wesley Professional, 2003.

[7] B. Caldwell, W. Chisholm, J. Slatin, and G. Vanderheiden. Web Content Accessibility Guidelines 2.0. W3C Working Draft, World Wide Web Consortium (W3C), 2006. http://www.w3.org/TR/WCAG20/.

[8] L. Carriço, N. Guimarães, C. Duarte, T. Chambel, and H. Simões. Spoken Books: Multimodal Interaction and Information Repurposing. In *Proceedings of HCII'2003, International Conference on Human-Computer Interaction*, pages 680–684, Crete, Greece, 2003.

[9] Crane Softwrights Ltd. Back of the Book Indexes using XSL-FO 1.0, 2006. Retreived from http://www.cranesoftwrights.com/resources/bbi/.

[10] D. Bulterman, G. Grassel, *et al.* Synchronized Multimedia Integration Language (SMIL 2.1). W3C Recommendation, World Wide Web Consortium (W3C), 2005. http://www.w3.org/TR/SMIL2.

[11] Specifications for the Digital Talking Book, 2005. http://www.daisy.org/z3986/2005/z3986-2005.html.

[12] S. DeRose, E. Maler, D. Orchard, and B. Trafford. XML Linking Language (XLink) Version 1.0. W3C Recommendation, World Wide Web Consortium (W3C), 2000. http://www.w3.org/TR/xlink/.

[13] C. Duarte, L. Carriço, N. Guimarães, and T. Chambel. Producing DTB from Audio Tapes. In *Proceedings of the 5th International Conference in Enterprise Information Systems, ICEIS'03*, 2003.

[14] J. Fink, A. Kobsa, and A. Nill. Towards a user-adapted information environment on the Web. In *Multimedia and Standardization 98*, 1998.

[15] D. Heckmann. *Ubiquitous User Modeling*. PhD thesis, Department of Computer Science, Saarland University, November 2005.

[16] D. Heckmann, T. Schwartz, B. Brandherm, M. Schmitz, and M. von Wilamowitz-Moellendorff. GUMO - the General User Model Ontology. In *Proceedings of the 10th International Conference on User Modeling*, pages 428–432, Edinburgh, UK, 2005. LNAI 3538: Springer, Berlin Heidelberg.

[17] Innovative Rehabilitation Technology. eClipseWriter. http://www.eclipsereader.com/productsECW.html.

[18] N. Kaplan and Y. Chisik. In the company of readers: the digital library book as "practiced place". In *JCDL '05: Proceedings of the 5th ACM/IEEE-CS joint conference on Digital libraries*, pages 235–243, Denver, CO, USA, 2005. ACM Press.

[19] M. Kay. XSL Transformations (XSLT) Version 2.0. W3c recommendation, World Wide Web Consortium (W3C), 2007. http://www.w3.org/TR/xslt20/.

[20] A. Kobsa. Generic User Modeling Systems. *User Modeling and User-Adapted Interaction: The Journal of Personalization Research*, 11(1-2):49–63, 2001.

[21] S. Little and L. Hardman. Cuypers Meets Users: Implementing a User Model Architecture for Multimedia Presentation Generation. Technical report, CWI - Centrum voor Wiskunde en Informatica, 2003.

[22] R. Lopes and L. Carriço. Building Reusable XML Pipelines with APP. In *XATA2006, XML: Aplicações e Tecnologias Associadas*, 2006. Available at http://vecpar.fe.up.pt/xata2006/papers/9.pdf.

[23] R. Lopes and L. Carriço. Automating XML Pipelines Through Rules. In *XATA2007, XML: Aplicações e Tecnologias Associadas*, 2007. Available at http://xata.fe.up.pt/2007/papers/13.pdf.

[24] R. Lopes, L. Carriço, and C. Duarte. DiTaBBu: Automating the Production of Time-based Hypermedia Content. In *WWW '06: Proceedings of the 15th international conference on World Wide Web*, pages 905–906, New York, NY, USA, 2006. ACM Press.

[25] R. Lopes, H. Simões, C. Duarte, and L. Carriço. Rich Digital Books for the Web. In *Web Information Systems and Technologies (WEBIST'2007)*, 2007.

[26] D. Newman, A. Patterson, and P. Schmitz. XHTML+SMIL Profile. W3c note, World Wide Web Consortium (W3C), 2002. http://www.w3.org/TR/XHTMLplusSMIL.

[27] I. J. Pitt and A. D. N. Edwards. Improving the usability of speech-based interfaces for blind users. In *Assets '96: Proceedings of the second annual ACM conference on Assistive technologies*, pages 124–130, Vancouver, British Columbia, Canada, 1996. ACM Press.

[28] P. Plessers, S. Casteleyn, Y. Yesilada, O. D. Troyer, R. Stevens, S. Harper, and C. Goble. Accessibility: a Web engineering approach. In *WWW '05: Proceedings of the 14th international conference on World Wide Web*, pages 353–362, New York, NY, USA, 2005. ACM Press.

[29] B. N. Schilit, G. Golovchinsky, and M. N. Price. Beyond paper: supporting active reading with free form digital ink annotations. In *CHI '98: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 249–256, Los Angeles, California, United States, 1998. ACM Press/Addison-Wesley Publishing Co.

[30] B. N. Schilit, M. N. Price, G. Golovchinsky, K. Tanaka, and C. C. Marshall. As We May Read: The Reading Appliance Revolution. *Computer*, 32(1):65–73, 1999.

[31] G. Schreiber and M. Dean. OWL Web Ontology Language Reference. http://www.w3.org/TR/owl-ref/, February 2004.

[32] P. Spyns, R. Meersman, and M. Jarrar. Data Modelling versus Ontology Engineering. *ACM SIGMOD Record*, 31(4), 2002.

[33] R. Stone, J. Dhiensa, and C. Machin. Profile-based Web Document Delivery. In *DocEng '06: Proceedings of the 2006 ACM symposium on Document Engineering*, pages 215–217, New York, NY, USA, 2006. ACM Press.

[34] O. Troyer, W. Goedefroy, and R. Meersman. UR-WSDM: Adding User Requirements Granularity to Model Web-based Information Systems. In *First Workshop on Hypermedia Development, Hypertext'98.*, 1998.

[35] O. D. Troyer and C. Leune. WSDM: A User Centered Design Method for Web Sites. *Computer Networks*,

30(1-7):85–94, Apr. 1998.

[36] J. van Ossenbruggen, J. Geurts, F. Cornelissen, L. Rutledge, and L. Hardman. Towards Second and Third Generation Web-Based Multimedia. *The Tenth International World Wide Web Conference*, pages 479–488, May 1-5 2001.

[37] W3C Device Independence Group. Composite Capability/Preference Profiles (CC/PP): Structure and Vocabularies 1.0. `http://www.w3.org/TR/CCPP-struct-vocab/`.

[38] N. Walsh, L. Muellner, and B. Stayton. *Docbook: The Definitive Guide*. O'Reilly, 2.0.8 edition, 2002.