Modeling Web Accessibility for Rich Document Production

RUI LOPES and LUIZ CARRICO

Faculty of Sciences, University of Lisbon, Lisbon, Portugal

This article presents a model-centered approach for rich document production and delivery to the accessible Web. A profile modeling task is proposed, where multidisciplinary teams can discuss users, devices, and usage scenarios, to grasp and synthesize rich scenarios for document delivery on the Web. A document production framework is presented, composed by different accessibility-aware components, affording the range of scenarios that might be modeled. By using this approach, documents are tailored to users in such a way that rich interaction capabilities are maintained, without sacrificing content accessibility.

KEYWORDS Web accessibility, profile modelling, document production, rich accessible documents.

INTRODUCTION

Today, Web accessibility is gaining momentum. Governments have started to legislate toward info-inclusion, resulting in increasing accessibility awareness. Guidelines are being defined to provide helpful cues for creating accessibility-friendly contents, such as WCAG (Web content accessibility guidelines)(Caldwell, Chisholm, Slatin, & Vanderheiden, 2006). However, awareness and guidelines are just the beginning of delivering highly interactive, yet accessible, documents on the Web. If fact, if we look into the real Web, we can see that accessibility is little more than providing overly simplified versions of contents. Consequently, users that rely on accessible contents have limited user experiences, when compared to those of the average user.
Accessibility can be influenced by very distinct factors. For example, those with physical disabilities, such as the blind or the deaf, require that document formats do not compromise information access (e.g., captioning pictures). Those with cognitive disabilities, such as the elderly, require special attention to navigation and interaction aspects within documents (e.g., fewer navigation items). On the other hand, the Mobile Web Initiative (2005) promises technological advances for accessing information anywhere. Therefore, device capabilities and limitations may pose significant accessibility problems even to nonimpaired users. Furthermore, when putting the user and devices in specific usage situations (e.g., outdoors), additional problems may arise when trying to access documents within a Web browser. Accessibility must be perceived nowadays as a crosscutting concern over users, devices, and usage situations.

At the same time, users are starting to demand richer interaction capabilities on the Web. Users want to interact with documents by annotating them, sharing opinions with others, etc. On the other hand, Web browsers are starting to provide the required technological support, such as being available in different devices and affording richer scripted behaviors. These capabilities should be exploited by document production frameworks, to deliver rich documents in any device and usage situation, without compromising users’ accessibility requirements on interactivity.

However, producing documents to the Web requires a multidisciplinary approach. Usability and accessibility experts, developers, graphical artists, managers, usability experts, and others should have a way to express and communicate these rich scenarios that does not compromise accessibility. Therefore, a high-level approach on understanding users, devices, and usage situations should be available to teams and, at the same time, tightly coupled with document production frameworks. To support this, this article presents a vision of rich accessible documents on the Web based on the definition of profiles for users, devices, and usage situations with a modeling methodology incorporating user-modeling ontologies. A supportive document production framework is also presented, affording different accessibility-aware tailoring components to process and deliver rich accessible documents to the Web.

**RICH ACCESSIBLE DOCUMENTS**

As of today, the information found on the Web can be regarded either as structured or semistructured. While structured data is perceived as stable tabular information, semistructured data do not fit this scenario. When data order is crucial, and structure is volatile and relatively unpredictable, we are faced with document-centered scenarios. This intrinsic characteristic poses several challenges on providing accessible, yet rich and interactive, documents to different kinds of users.
To understand these challenges, we must characterize the main five aspects of rich accessible documents: reading situations, enrichment, users, multimodality, and authoring.

Reading Situations

Users interact with documents through reading activities, which are greatly influenced by their goals. Reading a novel for entertainment purposes, and 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 categorization was proposed in Schilit et al., 1999. Two dimensions—nature of engagement and breadth of the activity—are used to categorize four conventional reading situations:

- **Passively reading a text**: Often associated with entertainment reading, passive reading occurs, for instance, when reading a novel;
- **Passively reading multiple texts**: frequently associated with keeping informed, e.g., when reading e-mail messages;
- **Actively reading a single text** is associated with learning tasks, as in studying a textbook;
- **Actively reading multiple texts**: commonly seen on research activities, where information is gathered from multiple sources.

While understanding the text is the common goal for all reading situations, each one poses different problems. Situations encompassing multiple texts entail the need to manage multiple documents and the difficulty of finding the needed information. Active reading (Adler & Doren, 1972) involves underlining, highlighting, and annotating, either on the text or separately, thus demanding some mechanism to manage annotations.

Digital documents and Web sites mitigate some of these problems. Digital documents offer the possibility to record, organize, and search annotations entered by the reader (Schilit, Golovchinsky, & Price, 1998). Web sites 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 Web site (such as a digital library) conveys a social sense to documents (Kaplan & Chisik, 2005).

Enrichment

Besides these capabilities, the document’s digital support opens up the possibility of enriching its content with supporting media (Carriço, Guimarães,
Duarte, Chambel, & Simões, 2003). 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 as an alternative to the visual presentation, similarly to digital talking books (DAISY Consortium, 2005). Copied with active reading perspectives, a document will start to be perceived as a hybrid between its intrinsic content and structure, and a reading platform.

Users

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

- Researchers, students, and other professionals 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. For instance, 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 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 customized visual components (e.g., increased font sizes and adequate color schemes); and
- Cognitively disabled users require adapted contents and interaction capabilities to achieve their goals in any reading situation (e.g., simplified content, reduced navigation tasks).

Multimodality

Inherent in these issues are the questions about output and input modalities, and how they fit into Web accessibility scenarios. 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., audio narration synchronized with visual
presentation). Consequently, the classification of the act of reading can be augmented according to users’ perceptions, to better adapt to their accessibility requirements, as follows:

- **Textual**: This is the traditional form of reading, where the document is presented in a textual format, suitable to various 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 centered on the visual interpretation of graphics (e.g., images, videos);
- **Aural**: Visually stressing situations will require an audio counterpart of a document (whether in the form of a speech track or real-time synthesized speech), to lower cognitive stress. Thus, reading will be perceived in an audio listening scenario.

In interacting with a rich accessible document, input modalities must also be selected appropriately, 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 focused on reading activities, not the opposite. Otherwise, an overwhelming cognitive effort from users will dismiss the benefits of rich accessible documents.

**Authoring**

Despite the fact that previous characteristics will be ultimately reflected on implementation details, the complexity of rich accessible documents requires a multidisciplinary approach at early conception stages of document-delivery environments. Therefore, a high-level discussion centered 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 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 im-
portantly, errors become easy to introduce within and between documents,
affecting the users’ overall reading experience. Adding up the complexity
and richness of the whole spectrum of scenarios described previously, de-
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 dis-
cuss users and usage situations at a high level and, at the same time, take
advantage of this knowledge of creating and configuring automated au-
thoring environments. This way, rich accessible documents will be always
tailored to users in a coherent way. Such a modeling task is presented in the
next section.

PROFILE MODELING

To get a better understanding of user, device, and usage situation charac-
terizations, the concept of profile modeling emerges. In a nutshell, a profile
reflects a document interaction scenario by grouping miscellaneous charac-
terizations in a systematic way. To cope with the broad range of scenarios
previously discussed, a profile modeling task should meet the following
requirements:

- **High-level perception**: Experts in different areas should easily grasp profile
  modeling. It should be easy for nondevelopers to model profiles with-
out having to learn low-level constructs such as scripting, expert shells,
etc. Moreover, a graphical presentation of profiles will be helpful 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
  emphasized (i.e., tailoring);
- **Direct applicability**: If profile modeling is excessively informal, it will only
  have the purpose of guiding developers on document production imple-
mentation tasks. Consequently, incoherence between models and their
reflection on production frameworks may yield critical accessibility prob-
lems 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 formalization, profiles can be integrated in adaptive Web platforms.

Based on these assumptions, a high-level modeling tool has been created with the Eclipse Modeling Framework (Budinsky, Steinberg, Merks, Ellersick, & Grose, 2003). Within this environment, a profile metamodel was specified, as depicted in Figure 1:

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 already gathered, e.g., through contextual design methodologies (Beyer & Holtzblatt, 1998). Each attribute will be a representation of a concept that characterizes 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 necessarily false within the profile’s domain, it is simply unknown.

With these basic constructs any number of profiles can be defined toward miscellaneous document-processing domains. However, as profile count increases, the probability of having profiles sharing concepts rises. Consequently, the profile metamodel introduces a specialization 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 specializations (i.e., inheriting attributes from more than one profile), allowing the modeling of more complex scenarios.

Going deeper on profile characterization through attribute definition, two approaches can be envisioned: informal vs. formal. The informal
approach 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, as previously explained, having a direct applicability is critical. Therefore, informal characterizations 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 characterizing profiles for users, devices, and usage scenarios. GUMO (general user model ontology)(Heckmann, Schwartz, Brandherm, Schmitz, & Wilamowitz-Moellendorff, 2005) provides a thorough research on user modeling in the form of a Web Ontology Language (OWL) (Schreiber & Dean, 2004) ontology, with a set of concepts ranging from abilities (e.g., AbilityToHear) to emotional states (e.g., Satisfied). On device characterization, CC/PP (W3C Device Independence Group, 2004)—Composite Capability/Preference Profiles—provides an RDF (resource description framework)(Beckett, 2004) framework to define different device profiles (named components). However, being just a framework, it does not provide the required richness of device characterization instances to be promptly used in profile modeling tasks. UbisWorld (Heckmann, 2005), an ontology that complements GUMO 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 (Spyns, Meersman, & Jarrar, 2002) can define other concepts to be used in profile modeling instances.

To illustrate the modeling 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, a 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:

This model defines the Desktop profile as a conjunction of ontology concepts that characterize 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 have 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, using a not() operator [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 in Figure 3, encompassing the usage situations described in the scenario.

Here, three usage situations provide a better understanding of the limitations they impose. The Desktop + Casual Reader profile states that the

![Figure 3](image3.png)

**Figure 3** Profile model with usage situations.
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 informa-
tion to transform documents toward 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 might become complex to cope with richer scenarios. This should
yield a set of good practices that should be taken into account when engi-
eering profiles for rich accessible documents on the Web. However, it is
not within the scope of this article to define such good practices. Next, a
production framework for rich accessible documents is presented.

PRODUCTION FRAMEWORK

After describing how profiles can be modeled toward Web accessibility sce-
narios, through the specification of user audiences, devices, and usage sit-
tuations, these should be exploited by document production frameworks.
Such frameworks process document structures toward some output format
for reading platforms (e.g., Web browser).

Therefore, a flexible production framework was redesigned and im-
plemented, based on previous results from automated rich document pro-
cessing (Lopes, Carriço, & Duarte, 2006; Lopes, Simões, Duarte, & Carriço,
2007). The key aspect of this framework centers on creating reusable and
accessibility-aware document processing components that can be recom-
bined. This allows the production of documents according to modeled pro-
files. The inheritance mechanism of profile modeling affords the necessary
cues to grasp which components may be reused across different production
configurations.

When redesigning the production framework, the main focus was to be
able to cope with different profiles, independent of the ontology concepts
used. Therefore, at the conceptual level, this framework was split into a
sequence of five concerns: content processing, structure repurposing, output
formatting, interaction, and presentation (as depicted in Figure 4). Each
concern aggregates a set of expandable processing components specialized
on different tasks (e.g., deleting subsections from a table of contents), which
may be used on runtime, according to the modeled profiles and respec-
tive framework configuration. This way, project managers can coordinate
their teams according to the processing components required to support the
FIGURE 4 Production framework overview.

profiles (e.g., teaming a usability expert with a developer to create a navigation scheme optimized for Mobile-based profiles).

Producing a tailored document on the framework must respect a specific processing flow, as follows: The framework receives on its input an initial content (e.g., any document), which is then processed by components along a specific configuration of the five concerns (in compliance with the modeled profiles). The resulting outcome is 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.

As stated before, rich accessible documents can be described through the composition of static media, dynamic media (i.e., audio, video), or even a mix of both. Despite the base medium of rich accessible documents, their meta-structure (e.g., chapter, section, paragraph, etc.) is highly stable. Therefore, DocBook (Walsh, Muellner, & Stayton, 2002) has been embraced by the production framework as the standard for describing documents and processing them with the framework’s components.

However, as DocBook by itself does not provide elements 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.

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

```
<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.
in the text vs. *two thousand and seven* in the audio). Appropriate descriptions were conceived also for video and images, in order to identify document excerpts within different media files.

However, having unrelated representations of the same content is not sufficient, especially as nontextual document sources do not have semantic information about the document’s structure (therefore decreasing significantly the interaction capabilities of a rich accessible document). Consequently, the production framework affords this type of relations based on XLink (DeRose, Maler, Orchard, & Trafford, 2000), according to the exemplificative syntax presented in Figure 6. Based on XLink’s extended link capabilities, more complex content relationship scenarios are also supported.

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 be reflected later on the resulting document, yielding severe accessibility problems (e.g., cognitive pressure induced by synchronization failure between audio and text presentation).

Regarding the implementation side of the production framework, we choose to use a different range of XML (extensible markup language) technologies, due to the nature of the document formats (as previously explained). This decision has reflected on the amount of code required to perform processing tasks, since XML technologies are particularly suited for document processing.

Consequently, XML pipeline architectures (Bruchez & Vernet, 2005; Lopes & Carriço, 2006) were used to layout the processing backend and framework configuration definitions. Accessibility-aware components were mostly implemented with XSLT (extensible stylesheet language transformations) 2.0 (Kay, 2007), coupled with special purpose Java classes where required.

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 have been implemented, and specialized toward the profile modeling scenario described earlier in this article. Nevertheless, existing components and new ones can be easily integrated and extended.
seamlessly in the framework, to support different rich accessible interaction scenarios.

Content Processing

The first concern within the document production framework defines a set of tasks centered on content processing. These tasks can be used on any framework configuration; therefore, they are independent from any profile model specified.

To reach a wider audience, there is the need to make available a big selection of books on different subjects. However, documents are stored in a wide range of formats (such as PDF, Microsoft Word, audio tracks, etc.) and must be normalized into the framework’s document formats (DocBook and auxiliary extensions). This way, each input document that is fed to the framework does not have accessibility barriers on its structure and content. Since manual conversion is error prone and cumbersome, the production of rich accessible documents with frameworks can only scale up if it is able to cope with already existing corpora automatically. Therefore, in order to accomplish this task, the production framework has an initial processing component based on XPR, XML Pipeline Rules (Lopes & Carriço, 2007). In a nutshell, XPR allows the specification of document format sensors and transformation rules to afford the required automatic normalization.

A particular subtask of this component is responsible for transforming multimedia content into the framework’s core document formats. To accomplish this, multimedia content reasoning is applied to the normalized documents and their multimedia counterparts, according to the multimedia content classification architecture proposed in Lopes et al. (2007). Reasoning is performed with feature extraction, indexing, and classification algorithms. In parallel, this subtask stores processed multimedia corpora in the multimedia repository, for subsequent content-enriching tasks. As an example, when speech-based audio tracks are fed to the framework, this subtask applies speech alignment algorithms to extract the relations between audio and textual counterparts of documents, using the mechanisms presented in Carriço et al. (2003). This way, the profiles centered on aural reading situations, such as Mobile + Casual Reader and Desktop + Blind, will have richer content navigation capabilities.

Structure Repurposing

The second concern about the production framework specifies repurposing components to be applied on document structures, such as extracting navigation paths, document simplification, and perform chunking/pagination on documents. Similarly to the content-processing concern, these components are independent from particular accessibility issues, and can be configured
to meet a specific accessibility feature that might have to be taken into account.

Based on the range of characteristics that may be synthesized by profile modeling, document navigation paths can be different from profile to profile. Nevertheless, algorithmically speaking, extracting and tailoring navigation paths can be performed by a single component. Extracting navigation paths can be performed multiple times in order to provide users with more than one single way to navigate through rich accessible documents, such as tables of contents, side note lists, or even indexes. This extraction is performed over the document’s DocBook structure (e.g., by seeking chapter, section, and similar DocBook elements), by extrapolating it into separate navigation documents according to the auxiliary extensions document formats. Another algorithm affords “back of the book index” construction (Crane Softwrights Ltd., 2006). Having all these alternate navigation paths opens the way for richer experiences from users on interacting with documents, without sacrificing accessibility concerns. To meet the needs of each profile, this component can be configured accordingly, such as extracting a shallow table of contents (e.g., chapter level) to be presented on Mobile profiles.

The next processing component is specialized on simplifying document structures. Documents targeted to the Web are traditionally read on desktop environments. Consequently, their creation is based on the assumption of an existing significant hardware performance. However, when thinking of computational limited devices (such as Web-enabled mobile phones), these performance 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. This way, the outcome of this component lightens the computational requirements to parse and present the documents. This component also affords essentiality tracks techniques (Atkinson, Dhiens, & Machin, 2006) that can be further applied to adapt content to users’ requirements. Once again, the configuration of this component affords the different perspectives that should be taken into account for each modeled profile.

Lastly, another component has been made available on this concern. Its task is to perform chunking and pagination over documents, based on different algorithms. While chunking is performed over well-defined document structures (e.g., chapter, section), pagination can be done according to criteria that split structures between pages (e.g., word count). This will allow for an adequate improvement on resource-limited usage scenarios such as Mobile-based profiles, or to convey a more physical sense of documents (as traditional Web documents are typically made to fit into just one page).

Output Format
In order to be parseable and presented by Web browsers and other Web-based reading platforms, documents have to be translated into known for-
mats at some point in the document processing flow. This task is handled by the Output Format concern within the production framework.

Different factors may influence which output format should be chosen to be delivered to users. While profiles synthesize user, device, and usage scenarios in a coherent form, each profile within the model may require a particular output format transformation. In the scenario presented earlier in this article, for instance, Desktop-based profiles can benefit from the transformation of documents into the XHTML+SMIL (synchronized multimedia integrated language) format (Newman, Patterson, & Smitz, 2002), supported by Internet Explorer. This format affords the integration of traditional HTML contents with native multimedia content, as well as advanced multimedia synchronization capabilities. This way, situations such as *aural reading* are afforded on traditional browsing experiences. Another output format option is the transformation of document contents into SMIL presentations (Bulterman et al., 2005), where the playback emphasis is put on multimedia enrichment. Consequently, *aural reading* mobile scenarios, such as the Mobile + Casual Reader profile, become less cumbersome to afford.

Due to the particular nature of output formats, i.e., their dependency from specific profile characteristics, it is impossible to provide output format completeness on an out-of-the-box processing component. Nevertheless, the developed component affords the transformation of documents into XHTML+SMIL and SMIL formats in a format as neutral as possible. This way, the example profiles were supported seamlessly. However, it is expected that little effort will be required to adapt the component to support other profiles, since both output formats can cope with a wide range of profile characteristics.

Another processing component has been developed for this concern, to provide support for active reading situations. To support it, documents are augmented with adequate features (e.g., highlighting, annotating, etc.) through attached scripts, links to trigger activities, etc. However, as the document has already been transformed into a particular output format, this component has to support the chosen format. It should be noted that document output formats might have inherent limitations, 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.

Interaction and Presentation

At this point, the document’s current state is still not ready to be delivered to a chosen profile. While it can certainly be presented on playback platforms, it will have significant drawbacks on user acceptance and usage, such as being inaccessible, or unappealing aesthetics. Therefore, interaction and presentation capabilities must be introduced, to deliver the document with high accessibility compliance and overall quality. Two concerns have
been specified for the document production framework, to meet this type of requirements.

On the interaction concern, different components are responsible for embedding input device-specific capabilities in the current document state, where each component is specialized on a particular input device. Currently, mouse, keyboard, and speech interaction capabilities have been implemented. Such capabilities can be, for instance, producing a vocabulary for speech recognition, limited to the set of words found in the document. This way, speech-recognition accuracy might be increased and used with high acceptance rates on speech-based document interaction (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 explicitly 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 typical discrete interaction afforded by keyboards.

In the same way as the previous component, this type of capability can only be supported if the chosen output format affords them. For instance, as Internet Explorer allows embedding ActiveX components, speech recognition can be introduced on documents without requiring native support of this input modality directly on the browser. On the other hand, when targeting the production framework toward pure SMIL output format (e.g., the Mobile profile), SMIL players do not feature adaptation of interaction mechanisms. As such capabilities are left to their enclosing runtime environments, the speech-recognition component should not be used in Mobile profiles. This is a practical example that demonstrates that profile-modeling discussions by interdisciplinary teams capture these problems and inconsistencies. This way, interaction alternatives can be further explored.

Finally, the last concern of the production framework encompasses a set of components that tailor the presentation of documents to users. Two types of requirements are met within this concern: branding and accessibility. The first directly relates to recommended editor aesthetics, to convey a consistent look among different documents, where a specific component processes documents accordingly. Afterwards, the accessibility presentation component further tweaks the current styling in consonance with the selected modeling profile. This way, even if an editor recommends an inaccessible styling, say, for a partially sighted user, it will still become fully accessible. It is worth mentioning that these components afford the textual, graphical, and aural reading situations (as defined earlier). 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’s selected browsing platform. Consequently, both components are responsible to instantiate these configurations into the selected output format.
The outcome of these two components and, in summary, of the presentation concern, is a final document version that has rich interaction capabilities and strong accessibility requirements compliance. This document is presented to the user on a Web-based playback platform according to the production framework configuration corresponding to the selected modeled profile.

In the next section, we present different experiments on configuring the production framework, based on the example profile modeled earlier.

RESULTS

The profile models presented earlier in this article served as the basis for the definition of different configurations for the production platform. The three situations (Desktop + Casual Reader, Desktop + Blind, and Mobile + Casual Reader) encompassed different requirements, resulting in the implementation of several processing components. Several of these components were reused across profiles (e.g., normalizing 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 toward just one profile (e.g., adding mouse interaction to XHTML+SMIL-capable browsers). Nevertheless, all these components could be used in the cases where usage situations were not known a priori (e.g., Casual Reader and Desktop profiles), enforcing coherence among all scenarios.

Regarding the Desktop + Casual 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 in Figure 7, the enriched document affords different navigation paths (table of contents and side notes) over fully synchronized visual and aural representations of the content, presented 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 led to selecting audio as the primary source of content delivery, coupled with a simple visual presentation mechanism, as presented in Figure 8.

Lastly, the Desktop + Blind profile resulted in 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 Inter-
net 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’ pre-
ferred input devices (as this issue is critical to reach a satisfaction level from
users).

**DISCUSSION**

As of today, accessibility on the Web is regarded (when it is regarded)
simply as following ad hoc good practices or, at most, WCAG specifications
(Caldwell et al., 2006). This typically results in providing simpler versions
of documents or Web sites, therefore resulting in 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 article 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 formal-
ized, 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 behaviors that cannot be synthesized 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., color 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.

RELATED WORK

The User Modeling research field has provided a significant effort on understanding users, and how supportive systems can exploit this information to adapt contents to users’ requirements and preferences. Such systems, like AVANTI (Fink, Kobsa, & Nill, 1998), are able to adapt content and navigation toward previously defined user models. However, these systems are typically oriented to adapt Web sites (i.e., data-oriented applications), therefore are unable to cope with semistructured documents. Moreover, these systems rely on expert shells to model users, as thoroughly dissected in Kobsa (2001). This requires high skills both on creating user models, and even on understanding them, contrasting with the high-level profile modeling methodology presented in this article.

A similar approach to profile modeling and document delivery to the Web is described in (Stone, Dhiensa, & Machin, 2006), 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, or usage scenarios. This decreases the acceptance of documents in ubiquitous environments. Moreover, the described framework does not cope with rich document 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 article. These technologies are mainly centered on the Digital Talking Book (DTB) specifications, therefore inheriting its benefits and problems. State-of-the-art systems, such as eClipseWriter (Innovative Rehabilitation Technology, n.d.), 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 robotized voices over longer periods of time. TTS (text-to-speech) technologies also do not cope with miscellaneous ambiguous contents (Pitt & Edwards, 1996), such as *IV* read as *i v* instead of *four*, a problem solved with the production platform presented in this article, through the use of speech-alignment technologies. Moreover, as none of these systems is based on user-modeling capabilities, the ability to
produce documents toward the rich usage scenarios presented in this article is dismissed.

A framework for producing enriched multimedia content for the Web has been proposed earlier (Ossenbruggen, Geurts, Cornelissen, Rutledge, & Hardman, 2001), where several transformation steps are applied in succession on multimedia documents, tailored to different usage situations. The system also encompasses a user model-centered transformation step (Little & Hardman, 2003) to adapt documents to miscellaneous user requirements. However, these mechanisms are used toward personalization scenarios (not accessibility). Moreover, being centered on multimedia, the system loses its focus on document structures, therefore removing significant 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, Web Site Design Methodology (WSDM) (Troyer & Leune, 1998), defines a user-centered approach for creating Web sites. It supports user modeling (Troyer, Goedefroy, & Meersman, 1998), albeit a fairly limited mechanism. More recently, this methodology has been extended to support accessibility-specific requirements (Plessers et al., 2005), where application specifications are automatically annotated for accessibility scenarios. However, this approach is highly centered on data models, thus not fitting document-based scenarios. Other limitations of WSDM, in the context of this article, relate to its inadequacy for aural reading scenarios, as it is mainly centered in textual information.

CONCLUDING REMARKS

This article 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 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) that 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 toward richer scenarios as the Web evolves, including:
• Define an ontology that covers accessibility scenarios (such as users, devices, and usage scenarios);
• Configure the production framework to automatically take advantage of modeled profiles;
• Afford 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).

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.

REFERENCES


