Multimodal Accessibility of Documents

Georgios Kouroupetroglou and Dimitrios Tsonos
National and Kapodistrian University of Athens
Greece

1. Introduction

Traditionally a "document" was considered as a textual record. Schamber (1996) defined document as a unit "*consisting of dynamic, flexible, nonlinear content, represented as a set of linked information items, stored in one or more physical media or networked sited; created and used by one or more individuals in the facilitation of some process or project*". Buckland (1997) tries to answer the question "what is a document?" through a discussion on how far the meaning of "document" can be pushed and which are the limits of "documentation".

The evolution of Computer Science and Information Technology created a new perspective for the word "document". The concept "electronic documents" can be differentiated from the "printed documents" having specific characteristics (Schamber, 1996): easily manipulabable, internally and externally linkable, readily transformable, inherently searchable, instantly transportable and infinitely replicable.

The term document used in this chapter refers to all kind of printed or electronic documentation the content of which is most text, like newspapers, books, journals, magazines, educational material, letters, brochures, leaflets, etc.

A document contains elements that arrange content in the page or even in the document itself. For example, the title of a chapter can be recognized by placing it on the top of the page and in larger font size than the body of text. Also, page indexing at the end of a document links the reader to specific part of the document.

The document functionality can be distinguished into browsing, searching, navigation and reading (Doermann et al., 1998). Additional concepts related to document's functionality are legibility, readability and aesthetics.

Legibility is a measure of easiness to distinguish one letter from another in a particular typeface and readability is a gauge of how easily words, phrases and blocks of copy can be read. These two measures were introduced as a comparison between printed and electronic documents on a computer screen (Mills & Weldon, 1987). Readability and legibility are closely related to typographic elements, typeface design and font/background color combinations (Hill & Scharff, 1997; Richard & Patrick, 2004; Eglin & Bres, 2003). Readability is more related to the overall and layout structure of a document (Holmqvist & Wartenberg, 2005; Holmberg, 2004; Kupper, 1989; Wartenberg & Holmqvist, 2005; Axner et al., 1997). Aesthetics of a document play a significant role during the reading process as well as reader's preferences (Porat et al., 2007; Harrington et al., 2004; Laarni, 2003).

"Dynamic" concepts like navigation and browsing are related to the interaction process with the reader. Navigation in electronic documents is a set of instructions to create an appropriate flow of a document for each type of devices (W3C, 2008a). The traditional way for navigation is by simply scrolling the window that presents the electronic document (like browsing in printed). By digitizing any printed document and transforming it into electronic format, new techniques derive that lead to a more sufficient and versatile navigation process combining navigation and browsing (W3C, 2008b; Czyzowicz, 2005; Cockburn et al., 2006). Thus, the functionality of a document can be distinguished into two tasks (Tsonos et al., 2007a):

- **Presentation task**, as an output; e.g. presenting the content and the information to the reader.
- **Navigation task**, as an input; e.g. performing actions by the reader, like searching or browsing for specific information.

Print disabilities prevent people from reading standard printed documents. They can be due to a visual, perceptual or physical disability which may be the result of vision impairment (blindness, low vision or dyschromatopsia), a learning disability (including dyslexia) or a disability (such as loss of dexterity) that prevents the physical holding of a book. Print-disabled are referred also as print-handicapped or read-disabled. Demographics for the print-disabled varied from 10% of the general population in Canada (IELA, 2008) to 17.5% in Australia (RPH, 2008).

People with print disabilities require printed documents in alternative formats, such as Braille, audio, large print or electronic text. They may also require assistive technology to meet their information needs. Braille displays provide in real-time text information into haptic modality. Screen magnifiers are software applications that help people with either low vision or dyschromatopsia to read documents. Text-to-Speech (TtS) is a common software technology that converts in real-time any electronic text into speech (Fellbaum & Kouroupetroglou, 2008). In most cases the text on a Graphical User Interface is detected by a software application named screen reader, which feeds actually the TtS system. TtS can be applied not only in Personal Computers, but also in Smart Mobile Phones and Personal Digital Assistants (PDAs). Automated Reading Devices (ARDs) are stand-alone machines that can convert printed or electronic text to audible speech using Text-to-Speech. They do not require to be connected to any other device, like a computer. ARDs have been designed for use by individuals who are print-disabled. Depending on the type of the source material, there are two main classes of ARDs (Freitas & Kouroupetroglou, 2008):

- Printed-Text (PT) ARDs,
- Electronic-Text (ET) ARDs.

Most of the current Text-to-Speech systems do not include effective provision of the semantics and the cognitive aspects of the visual (such as font and typesettings) and non-visual document elements. Recently, there was an effort towards Document-to-Audio (DtA) synthesis, which essentially constitutes the next generation of the Text-to-Speech systems, supporting the extraction of the semantics of document metadata (Fourli-Kartsouni et al., 2007) and the efficient acoustic representation of both text formatting (Xydas & Kouroupetroglou, 2001a; 2001b; Xydas et al., 2003; 2005) and tables (Spiliotopoulos et al., 2005a; 2005b) through modelling the parameters of the synthesised speech signal.

According to Stephanidis (2001), **accessibility** concerns the provision and maintenance of access by disabled and elderly people to encoded information and interpersonal communication, through appropriate interaction with computer-based applications and telematic services. W3C (2008e) determines Web Accessibility as means that people with disabilities can perceive, understand, navigate, and interact with the Web and its content.

The last two decades there were many efforts in the domain of the accessibility of documents. Some of them deal with the web content accessibility for visually impaired users (Kouroupetroglou et al., 2007; Harper & Yesilada, 2007; Chen et al., 2006; Rosmaita, 2006). Bigham et al. (2006) studied how images can be accessible and Saito et al. (2006) proposes a method to transform existing Flash content into XML structures. Edwards et al. (2006) tried to create mathematics accessible to blind students and Francioni & Smith (2002) proposed a framework for the accessibility of computer science lessons for visually impaired students. Tsonos & Kouroupetroglou (2008) proposed recently a Design-for-All approach for accessible documents on the board and during the presentations in the classroom.

Multimodal interaction with documents is considered the execution of the presentation and navigation tasks according to reader's preferences in one of three modalities, visual, acoustic and haptic or in any preferable combination. Guillon et al. (2004) proposed an integrated publishing procedure for accessible multimodal documents based on DAISY 3.0 (DAISY, 2008). World Wide Web Consortium (W3C, 2008c) proposes guidelines for the multimodal interaction (W3C, 2008b). Multimodal Accessibility should support any device: thin clients (devices with little processing power or capabilities that can be used to capture user input - microphone, touch display, stylus, etc. - as well as nonuser input, such as GPS), thick clients (devices such as a PDA or notebook) and medium clients (devices with some degree of interpretation) (Mikhalenko, 2008). Besides the above efforts, there is still a challenging question for both the navigation and the presentation tasks: "Are documents accessible by anyone?"

In this chapter we first present an integrated architecture on how a document is structured. Then, the existing international standards and guidelines for creating accessible documents are given. Based on these, we propose in the following section a novel holistic XML-based system for the real time production, presentation and navigation of multimodal accessible documents.

2. Document Architecture

The way a document is composed and presented either on paper or on screen, refers to the term "document architecture" (Peels et al., 1985). In Figure 1 we propose a general document architecture, which constitutes an extension of the basic ITU/ISO model (ITU, 1993; ISO, 1989). One can identify two different but complementary views of a specific document:

- *Logical view*: associates content with architectural elements such as headings, titles /subtitles, chapters, paragraphs, tables, lists, footnotes and appendices.
- *Layout view*: associates content with architectural elements relating to the arrangement on pages and areas within pages, such as columns and margins.

Typography essentially includes font (type, size, color, background color, etc.) and typesetting (such as bold, italics, underline). It can be applied to both the logical and the layout view of a document. For example, the title of a chapter (Logical view) can be in bold or in larger font size than the body of the text. The vertical space in a text block, called leading (Layout view) can be affected by the font type. Furthermore, typography can be applied to the body of the text directly, e.g. a word in bold is used to indicate either emphasis or the introduction of a new term. "Typography exists to honour content" (Bringhurst, 1996).

There is an analogy of the above terminology with the classification of document elements proposed by Tsonos et al. (2007a):

- Text Formatting ↔ Typography

- Text Structure ↔ Logical View
- Text Layout ↔ Layout View

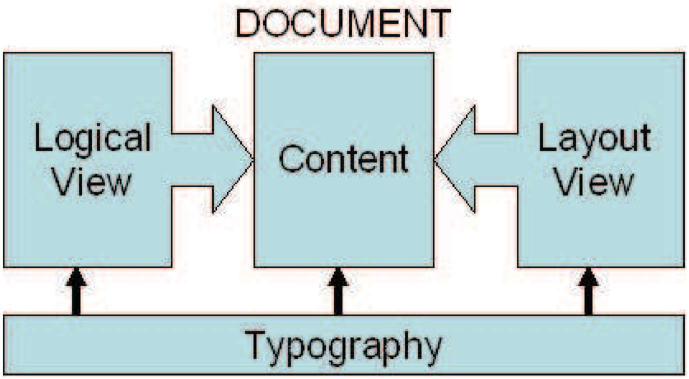


Figure 1. The general model of the document architecture

According to Tsonos et al. (2007a) the content of a document includes also non-textual elements, such as figures, drawing etc.
The document elements described above affect the readability, the legibility, as well as the aesthetics of a document. Axner et al. (1997) found that the three-column text format is easier to be read than single-column. Hall & Hanna (2004) examined the impact of text-background color combinations on readability and aesthetics providing the following results:

- colors with greater contrast ratio lead to greater readability,
- preferred colors (e.g. blues and chromatic colors) lead to higher ratings of aesthetic quality.

Laarni (2003) investigated the effects of color, font type and typesetting on user preferences. He concluded that: i) the most readable combinations are: i.1) plain Times New Roman black color font on white background, i.2) italized Arial white font on blue background and i.3) plain Arial white font on black background and ii) the less readable are the combinations of red font on green background. Also he examined their impact of color on document aesthetics (e.g. combinations of red font on green background were rated as the most unpleasant and black on white were considered the least arousing).

3. Standards and Guidelines of Accessible Documents

This section is intended to give a brief overview of the available guidelines and standards that are direct related with the accessibility of documents.

3.1 Web Accessibility Guidelines

Web Content Accessibility Guidelines (W3C, 2008d) are part of Web Accessibility Guidelines provided by Web Accessibility Initiative (W3C, 2008e). The scope of these guidelines is to create accessible web content to people with disabilities.
The guidelines concern all the Web content developers (page authors and site designers) as well as the developers of authoring tools. Following these guidelines the Web content

becomes device / software independent, thus it makes it accessible in any of three modalities. The guidelines address not only people with permentant impairements but also those with temporal or transient disabilities (e.g., a person that uses a mobile device and drives a vehicle in a noisy environment).

These guidelines help people to access the Web more quickly. Also, the developers are encouraged to use multimedia content (like images, videos, sounds etc.) but in a manner that is properly accessible.

The content of a web page can be presented by a *text* or *non-text equivalent*. These two terms are comprehensible using the following examples:

1. Suppose there is an image in a web page. A non-sighted user is unable to view the image. A description (*text equivalent*) of the image can make the image accessible to the user. The description can differ according to author's intentions, by simply describing the image that is supplementary to the main content-text of the web page or to guide the user to click the image for navigation purposes. Thus, two users with different needs can access the image using the same kind of web browser, sighted user can view the image and non-sighted can hear the description using synthesized speech through a Text-to-Speech or "read" the description through a Braille display.
2. Other multimedia features in web page are the pre-recorded speech (e.g. a welcome message in a site). A deaf user is unable to hear the welcome message. A text equivalent of the audio file can be accessed by the user and read the welcome message.
3. In the previous example, a deaf user can alternatively "hear" the description or the message using a video stream or a virtual agent that translates the description into sign language. This is the *non-textual* equivalents of text.

The WAI guidelines address two general goals: ensuring graceful transformation and making content understandable and navigable (W3C, 2008d). Web Content Accessibility Guidelines:

- Provide equivalent alternatives to auditory and visual content.
- Don't rely on color alone.
- Use markup and style sheets and do so properly.
- Clarify natural language usage.
- Create tables that transform gracefully.
- Ensure that pages featuring new technologies transform gracefully.
- Ensure user control of time-sensitive content changes.
- Ensure direct accessibility of embedded user interfaces.
- Design for device-independence.
- Use interims solutions.
- Use W3C technologies and guidelines.
- Provide context and orientation information.
- Provide clear navigation mechanisms.
- Ensure that documents are clear and simple.

3.2 Open Document Format Accessibility

The OpenDocument Format (ODF) is an open XML-based document file format for office applications to be used for documents containing text, spreadsheets, charts, and graphical elements. The file format makes transformations to other formats simple by leveraging and reusing existing standards wherever possible (ODF, 2008a, OASIS, 2008). The creation and support of this standard imposes the possibility for the implementation of new applications and the backward compatibility of the traditional office applications.

The ODF schema provides high-level information suitable for editing documents. It defines suitable XML structures for office documents and is friendly to transformations using XSLT or similar XML-based tools.

Under the guidelines and support of ODF, the Open Document specification for accessibility has been created. The specification intends to discover and improve accessibility issues and enhance the creation, reading and editing process of office documents for people with disabilities (ODF, 2008b). The Open Document Format comprises much structural and semantic information that is needed for the proper access to this information by people with disabilities.

Open Document accessibility subcommittee categorizes accessibility into three types of access: direct access, mediated access and indirect access (ODF, 2008b).

Combining the use of computer - assistive technologies and Open Document Accessibility specifications, a disabled user can have direct access to the content of a document.

The types of disabilities supported by ODF – Accessibility are:

- Minor vision impairments.
- Major vision impairments.
- Near or total blindness.
- Minor physical impairments.
- Major physical impairments without speech recognition.
- Major physical impairments with speech recognition.
- Hearing impairments.
- Cognitive impairments.

3.3 Math Markup Language

W3C through MathML (W3C, 2008f) proposes recomendations for the production and representation of mathematics in XML. Traditionally, the representation of mathematical expressions and scientific notation in electronic documents were implemented using pictures and images and the caption of the image as a description of the image (even in HTML). But, these figures are not accessible by, e.g. a blind student. This is because, a screen reader is not able to “read” the image but only the description provided by the caption or any metadata accompanying the image.

MathML is a low-level specification for describing mathematics as a basis for machine to machine communication. It provides a much needed foundation for the inclusion of mathematical expressions in Web pages (W3C, 2008f)

MathML has been designed with the following goals (W3C, 2008g):

- Encode mathematical material suitable for teaching and scientific communication at all educational levels.
- Encode both mathematical notation and mathematical meaning.
- Facilitate conversion to and from other mathematical formats, both presentational and semantic. Output formats should support: graphical displays, speech synthesizers, input for computer algebra systems, other mathematics typesetting languages, such as TEX, plain text displays, e.g. VT100 emulators, print media, including Braille.

It is recognized that conversion to and from other notational systems or media may involve loss of information (W3C, 2008g).

MathML combines not only visual representation tasks but also the meaning - semantics (like "divide", "times", "power of") of mathematical notations. For example, embedded mathematical expression using MathML in web pages can be viewed as normal web pages (visual representation task). Blind individuals can use a screen reader along with a Text-to-Speech system to hear the description of the same mathematic expression using MathML (semantic representation).
An application, basically for the acoustic rendering of the scientific notations using MathML, is the free available MathPlayer plugin (Soiffer, 2005; MathPlayer, 2008) for the web browsers.

3.4 Braille Markup Language

A new and promising standard for haptic representation of documents' content in Braille is the Markup Language (BrailleML). BrailleML is an effort towards the standardization of Braille documents. Masanori et al., (2007) proposes the automated conversion of ODF documents into Braille Documents using the BrailleML, which is an XML language described in the XML Schema.

3.5 Scalar Vector Graphics

Scalar Vector Graphics (SVG) is a language for describing two-dimensional graphics and graphical applications in XML (SVG, 2008). SVGs offer a number of features to make graphics on the Web more accessible, to a wider group of users. Users who benefit include; those with low vision, color blind or blind users, and users of assistive technologies. A number of these SVG features can also increase usability of content for many users without disabilities, such as users of PDAs, mobile phones or other non-traditional Web access devices (SVG, 2000).
Disabled users are provided with many accessibility features by the SVG specification. SVG images are scalable - they can be zoomed and resized by the reader as needed. Scaling can help users with low vision and users of some assistive technologies (e.g., tactile graphic devices, which typically have low resolution).

3.6 Ink Markup Language

InkML is an XML data format for representing digital ink data that is input with an electronic pen or stylus as part of a multimodal system. Ink markup provides a format for:

- transferring digital ink data between devices and software components,
- storing hand-input traces for: Handwriting recognition (including text, mathematics, chemistry), Signature verification, Gesture interpretation.

The InkML specification is designed by the ink subgroup of the Multimodal Interaction Working Group of the W3C (InkML, 2006). The InkML requirements can be divided into two categories (InkML, 2003):

- primitive elements, which represent low-level information about digital ink (like device and screen context characteristics and pen traces),
- application-specific elements, which characterize meta-information about the ink (a group of traces that belong to a field in a form).

3.7 DAISY/NISO standard

The DAISY Consortium (DAISY, 2008) is the official non-profit maintenance agency for the DAISY/NISO standard (officially known as ANSI/NISO Z39.86) (ANSI/NISO, 2008) provided by the National Information Standards Organization NISO. This standard defines the format and the content of the electronic file set that comprises a digital talking book (DTB) and establishes a limited set of requirements for DTB playback devices. This standard specifies the guidelines for the production and presentation of Digital Talking Books (DTBs) for print-disabled readers (blind, visually impaired, physically disabled and those with learning disabilities).

DAISY/NISO standard provides specifications for the format of DTB files (production of DTBs) and also sets the specifications for the DTB playback devices:

- Player performance related to file requirements.
- Player behaviour in areas defined in user requirements.

A Digital Talking Book (DTB) is: a collection of electronic files arranged to present information to the target population via alternative media, namely, human or synthetic speech, refreshable Braille, or visual display, e.g., large print (ANSI/NISO, 2008).

The files that comprise a DTB can be divided into the following ten categories:

- Package File.
- Textual Content File.
- Audio Files.
- Image Files.
- Synchronization Files.
- Navigation Control File.
- Bookmark/Highlight File.
- Resource File.
- Distribution Information File.
- Presentation Styles.

The merge of these files' functionalities and the creation of a DTB according to DAISY/NISO standard incorporates many features for either Navigation or Presentation tasks during the reading process. A few features are:

- In Navigation task: rapid - flexible navigation, bookmarking – highlighting, keyword searching.
- In Presentation task: user control over the presentation of selected items (e.g., footnotes, page numbers, etc.).

The navigation features provided by DTB include: Fast Forward and Fast Reverse, Reading at Variable Speed, Notes, Cross Reference Access, Index Navigation, Bookmarks, Highlighting, Excerpt Capability, Searching, Spell-Out Capability, Text Attributes and Punctuation, Tables, Nested Lists, Text Elements, Skipping User-Selected Text Elements, Location Information, Summary and Reporting Information, Science and Mathematics.

These features enable the presentation and navigation in DTB either in visual, acoustic or haptic modality. According to reader's disabilities and needs, the document can be presented and accessed in multiple ways and combinations of the three modalities.

Following this standard one can implement a DTB player with a variety of capabilities and functionalities. A DTB player can vary from a portable device, which simply "reads" to the user the content of the book using synthetic speech supporting basic navigation features like fast forward and reverse, reading at variable speed, to a more efficient PC-based player, supporting all modalities and all navigation features provided by the standard.

4. An XML-based system for Multimodal Accessibility of Documents

The discussion on accessibility of documents imposes questions on how the documents can be accessed either in visual, acoustic or haptic modality. Studies, such as (Edwards et al., 2006; Raman, 1992; Xydas & Kouroupetroglou, 2001b), are trying to create accessible documents in the acoustic modality using speech synthesis or by combining earcons (Kramer, 1994; Brewster et al. 1996; Mynatt, 1994), auditory icons (Gorny, 2000) and 3D sounds (Djennane, 2003) for the auditory. Recently a rather ad-hoc approach for multimodal accessibility of mainly TEX formatted technical documents was introduced by Power (2008). In this section we present a novel integrated XML-based system for the real-time production, presentation and navigation of multimodal accessible documents by conforming to the guidelines and standards discussed in section 3. This approach includes a unified methodology for the multimodal rendering of text formatting, text structure, text layout and non-textual elements.

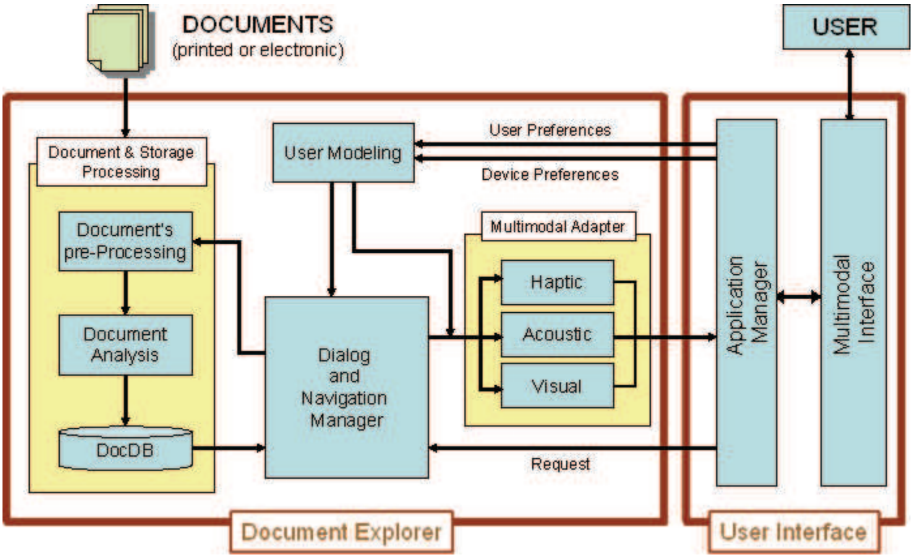


Figure 2. The XML-based system for multimodal accessibility of documents

The overall proposed architecture consists of two main parts (Figure 2): the *User Interface* and the *Document Explorer*. The later is responsible for document analysis during the production process and the exploration in the documents (in the navigation task). User Interface is responsible for the multimodal interaction with the documents. It collects the user preferences and the device requests, as well as the navigation commands, and executes the presentation task. These parts are implemented by two different modules following the Client – Server model. The Document Explorer can be hosted on a powerful server machine due to the resource demanding tasks that performs. In contradiction, the User Interface can be hosted on any common computer (e.g. a personal computer, a PDA, mobile smart phone). This kind of implementation fulfills the Web Content Accessibility Guidelines, in order to be device and software independent. The implementation and the communication between the modules are XML-based.

The presentation that follows is given according to the:
- Production task,
- Navigation task,
- Presentation task.

4.1 Production of Multimodal Accessible Documents

The *Documents' Processing and Storage* module is responsible for the production of accessible documents. It parses a document and then creates the output according to DAISY/NISO standard. The output is stored in a database (DocDB), so the Dialog and Navigation Manager module can have quick access to the content.

Document pre-processing

Figure 3 illustrates the document's pre-processing module. It can handle either printed or electronic documents. The *printed* documents are scanned and pass an Optical Character Recognition (OCR) software application. The Markup Normalization Module parses the digitized or the electronic documents in order the file format conforms to the DAISY/NISO specification. The output file (namely the Document-ML) includes: the content of the document, text formatting, text structural, text layout and non-textual elements (as described in section 2).

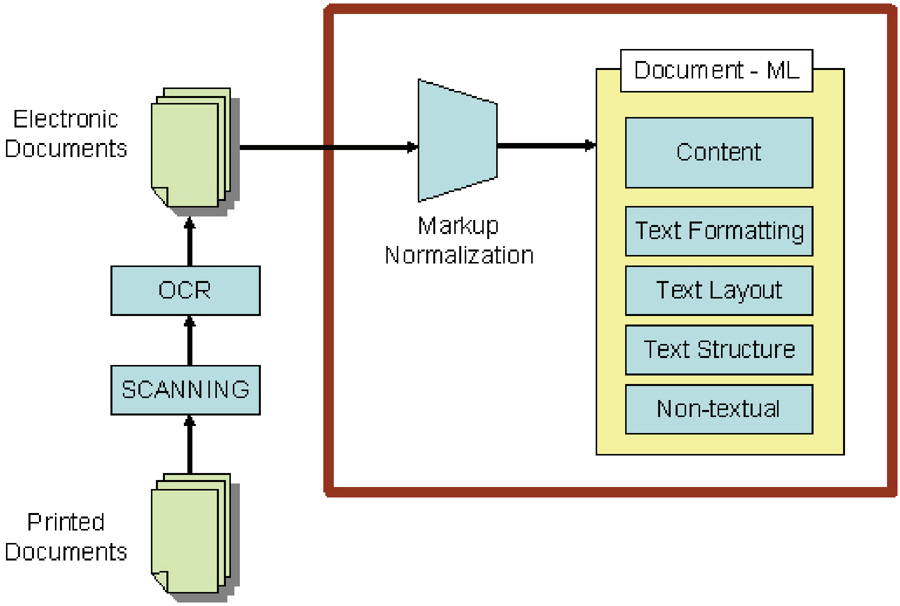


Figure 3. Documents' Pre-processing Module

To serve the multimodal accessibility requirements, the architecture tags the document at the:
- *Semantic* layer,
- *Emotional* layer.

Both these layers are free of presentation details. The Semantic layer aims to record the reader's understanding of the document. The later layer encodes the emotional state of the reader during the reading process. The output data of these two layers can be transformed to any modality. In the following paragraphs we present the way the original visual stimulus affects the multimodal presentation, via these layers, emphasizing the acoustic modality.

Semantic layer

Recently there was an attempt to produce an automatic extraction system of semantic information based only on the document layout, without the use of natural language processing (Fourli-Kartsouni et al, 2007). However, there are several studies on the automatic identification of logical structure of documents e.g. (Conway, 1993; Yamashita et al., 1991; Derrien-Peden, 1991; Krishnamoorthy et al., 1993). Most traditional approaches in this field have employed deterministic methods (decision trees, formal grammars) (Mao et al., 2003; Tsujimoto & Asada, 1990; Derrien-Peden, 1991), which may suffer from poor performance due to noise and uncertainty. In addition, such approaches create models which are not flexible to domain changes and cannot easily evolve in the presence of new evidence. In order to overcome such limitations, Fourli-Kartsouni et al. (2007) employed a probabilistic approach based on Bayesian networks trained on a series of labelled documents. Bayesian networks offer a significant tolerance to noise and uncertainty, they can capture the underlying class structure residing in the data and they can be trained on examples, thus adapting to existing and future evidence. It can learn the mapping between text's formatting, structure, layout and logic elements. The mapping rules (which in some cases is 1:N) can be derived from a series of experiments, (e.g. bold can be mapped as emphasis but also as strong emphasis).

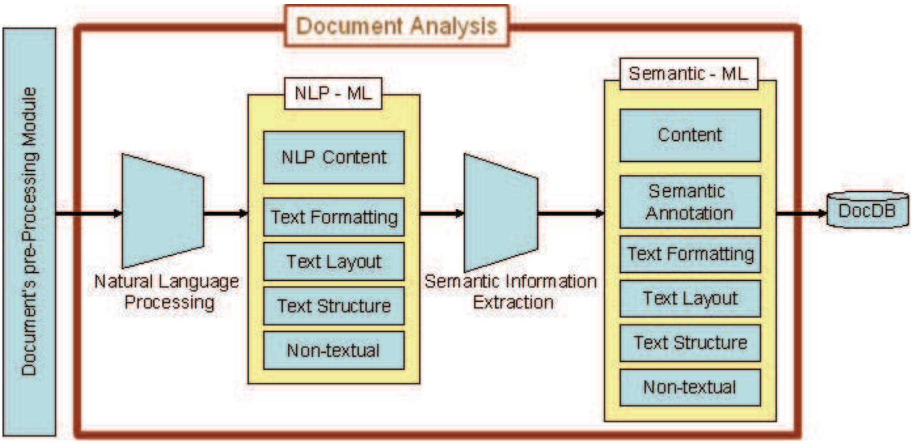


Figure 4. Document Analysis module: The Semantic Information Extraction approach

The Document Analysis Module (Figure 4) adds semantic annotation tags to documents using the methodology proposed by Fourli–Kartsouni et al. (2007). The module produces an XML-file (Semantic-ML) including all the document elements described in section 2 and stores the output on the database (DocDB).

Emotional layer

Many studies in the field of Human Computer Interaction focus on the user's emotional response during the interaction (Kärkkäinen & Laarni, 2002; Humaine Portal, 2008;

Dormann, 2003). The document elements affect directly the reader's emotions, emotional state and the readability of the document. Multiple combinations of colors (Birren, 1984), font size, type and style in a document affects the emotional state (Laarni, 2003; Sánchez et al., 2005; Sánchez et al., 2006; Tsonos et al., 2008) and the readability of the document (Laarni, 2003; Saari et al., 2004) not only in printed but also in electronic documents (Larson, 2007).

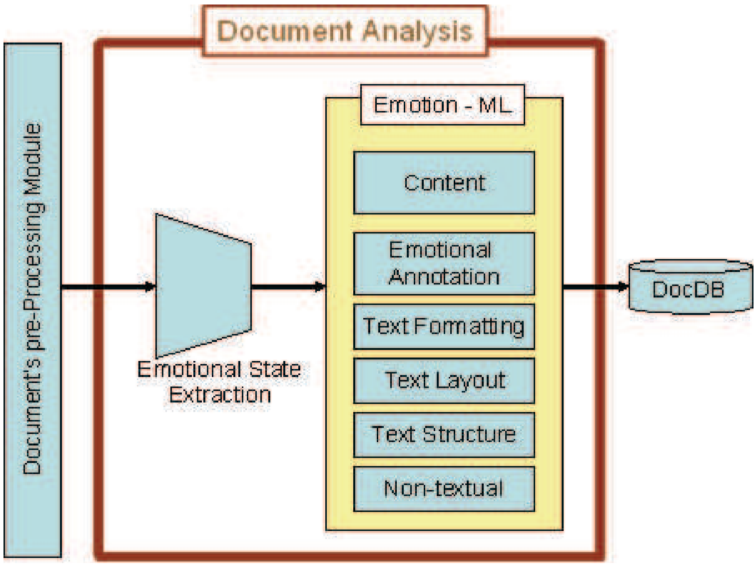


Figure 5. Document analysis module: The Reader's Emotional State Extraction approach

The Document Analysis module (Figure 5), using the automated readers' emotional state extraction (Tsonos et al., 2007b, Tsonos et al., 2008), implements the mapping of document's elements into variations of the emotional states Pleasure, Arousal and Dominance (P.A.D.) or specific emotions. The module produces an XML file (Emotion-ML) with emotional state annotation.

4.2 Navigation in Multimodal Accessible Documents

The Dialog and Navigation Manager handles the navigation tasks and forwards the content to Multimodal Adapter. The manager supports all the navigational functionalities suggested by the DAISY/NISO standard using the Navigation Control File (this file describes the hierarchy of the document).

For the implementation of the manager it is optimum to use an adaptive, agent-based dialog system as proposed in (Frink, 1999; Hjalmarsson, 2005) to support:

- System Use, taking over parts of routine tasks, adapting the interface, giving advice of system use, controlling a dialogue,
- Information Acquisition, helping users to find information, tailoring information presentation.

Such an agent-based dialog system allows every module to interact and is capable of reasoning. It is a flexible allowing the user to have full control of the dialog (McTear, 2002) and supports multimodal dialog input (Turunen et al., 2005).

User preferences and device requests are processed by the User Modeling module and are handled by the Dialog and Navigation Manager. *User Modeling* module creates the user's profile according to her/his interaction and device requests. As an example, the user is able to select the desired modality for the presentation of document's content according to her/his needs or the way a command is given to the system, along with the requirements of the user's device. The user profile is a collection of actions and preferences acquired by the interaction with the Application Manager. Such can be the user's needs, the interaction modality, the navigational factors, environmental factors etc. The produced XML file (User-ML) is forwaded to the Dialog and Navigation Manager. Thus, depending on the user preferences, the dialog manager handles the way the content will be treated and delivered in each modality (acoustic, visual, haptic) through the Acoustic Adapter or/and the Visual Adapter and/or the Haptic Adapter.

4.3 Presentation of Multimodal Accessible Documents

User Interface includes two parts: the *Multimodal Interface* and the *Application Manager* (Figure 2). Multimodal Interface is responsible for the user's interaction with the documents supporting the three interaction modalities visual, acoustic and haptic. The User Interface supports the:

- input tasks (operation or navigation commands): accomplished by using either input devices (e.g. keyboard, mouse, buttons) or input applications (e.g. Automatic Speech Recognition),
- output tasks: The user receives the requested content or system prompts in sequential and hierarchical navigation.

The functionality of the above tasks can be changed according to user's preferences.

The Application Manager collects and handles the device requests and the user preferences provided by the Multimodal Interface. The output of the Application Manager is used by the: a) User Modeling module and b) Dialog and Navigation Manager (navigation or application commands).

The Dialog and Navigation Manager feeds the Multimodal Adapter with the document that should be presented in Semantic-ML or Emotion-ML. The corresponding adapter is triggered and produces an output data according to the user preferences and the device requests. The output is handled by the User Interface for the content presentation in visual, acoustic or haptic modality.

Focusing on the acoustic modality the mapping is obtained using a Document-to-Audio (DtA) platform (Xydas & Kouroupetroglou, 2001b). In the acoustic adapter, the Adaptation module (Figure 6) combines information about user preferences (User Modeling) and the rules for the acoustic mapping (CAD script) so the result can be used by the e-TSA Composer. The Cluster Auditory Definition (CAD) scripts provide the mapping rules to DtA platform. The rules, that are used to describe the relation of document's elements and acoustic representation, depend on the methodology that is followed, using the semantic (Xydas et al., 2003, Spiliotopoulos et al., 2005) or the emotional approach utilizing Expressive Speech Synthesis (Campbell et al., 2006; Pitrelli et al., 2006; Schröder, 2006).

The DtA platform offers greater priority to user preferences than the default CAD rules. For example, the user might need to hear faster the content of the book, but some elements should be read slower. The Adaptation Module will give higher priority to the rules provided by User Modelling. Using the DtA platform, documents are mapped into specific acoustic elements which are realised by the auditory synthesizer (the output format can be e.g. MPEG4, SMIL or WAV).

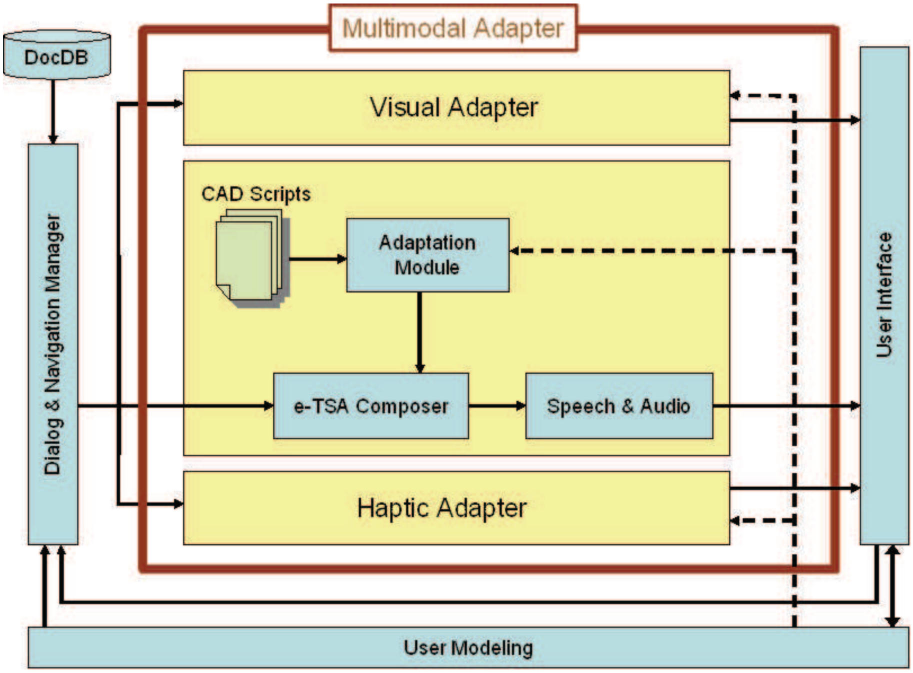


Figure 6. Multimodal Representation

5. Conclusions

To ensure that people with print disability are able equally to participate in society, it is crucial to develop more effective ways for the accessibility of both printed and electronic documents.
In this chapter we have first presented an integrated architecture on how a document is structured. Then, the existing international standards and guidelines for creating accessible documents were briefly analyzed. Based on these, we have proposed a novel holistic XML-based system for the real time production, presentation and navigation of multimodal accessible documents.

6. Acknowledgements

The work described in this chapter has been funded by the European Social Fund and Hellenic National Resources under the HOMER project of the Programme PENED, Greek General Secretariat of Research and Technology.

7. References

ANSI/NISO (2008). National Information Standards Organization, The DAISY standard, *http://www.niso.org/workrooms/daisy/*

Axner, E.; Strom B.; Linde-Forsberg, C.; Dyson, M. C. & Kipping, G. J. (1997). The legibility of screen formats: Are three columns better than one? *Computers & Graphics*, Vol. 21, No. 6, December 1997, pp. 703-712, ISSN: 0097-8493

Bigham, J. P.; Kaminsky, R. S.; Ladner, R. E.; Danielsson, O. M. & Hempton, G. L. (2006). WebInSight: making web images accessible, *Proceedings of the 8th international ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '06)*, pp. 181-188, ISBN: 1-59593-290-9, Portland, Oregon, USA, 23-25 October 2006, ACM, New York

Birren, F. (1984). *Color & Human Response: Aspects of Light and Color. Bearing on the Reactions of Living Things and the Welfare of Human Beings*, J. Wiley & Sons Inc, ISBN: 978-0-471-28864-0, New York

Brewster, S. A.; Räty, V. & Kortekangas, A. (1996). Earcons as a Method of Providing Navigational Cues in a Menu Hierarchy, *Proceedings of HCI on People and Computers XI*, pp. 169-183, ISBN: 3-540-76069-5, M. A. Sasse, J. Cunningham, and R. L. Winder, Eds. Springer-Verlag, London

Bringhurst, R. (1996). *The elements of typographic style,* 2d ed. Hartley and Marks, ISBN: 0881791326, Vancouver

Buckland, M. K. (1997). What is a "document"? *Journal of the American Society for Information Science,* Vol. 48, No. 9, September 1997, pp. 804-809, ISSN: 0002-8231

Campbell, N.; Hamza, W.; Hoge, H.; Tao, J. & Bailly, E. (2006). Special Section on Expressive Speech Synthesis. *IEEE Transactions on Audio, Speech and Language Processing*, Vol. 14, No. 4, July 2006, pp. 1097-1098, ISSN: 1558-7916

Chen, X.; Tremaine, M.; Lutz, R.; Chung, J. & Lacsina, P. (2006). AudioBrowser: a mobile browsable information access for the visually impaired. *Universal Access in the Information Society*, Vol. 5, No. 1, June 2006, pp. 4-22, ISSN: 1615-5297

Cockburn, A.; Gutwin, C. & Alexander, J. (2006). Faster document navigation with space-filling thumbnails, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '06)*, pp. 1-10, ISBN: 1-59593-372-7, Montréal, Québec, Canada, 22-17 April 2006, ACM, New York

Conway, A. (1993). Page grammars and page parsing: a syntactic approach to document layout recognition, *Proceedings of the Second International Conference on Document Analysis and Recognition*, pp. 761-764, ISBN: 0-8186-4960-7, Tsukuba Science City, Japan, 20 – 22 October 1993

Czyzowicz, M. (2005). Intelligent Navigation in Documents Sets Comparative Study, *Proceedings of the International Intelligent Information Processing and Web Mining Conference (IIPWM' 05),* pp. 421-425, ISBN: 978-3-540-25056-2, Gdansk, Poland, 13-16 June 2005, Springer, New York

DAISY Consortium (2008), *www.daisy.org*

Derrien-Peden, D. (1991). Frame-based system for macro-typographical structure analysis in scientific papers, *Proceedings of International Conference on Document Analysis and Recognition (ICDAR91)*, pp. 311-319, Saint-Malo, France, September 1991

Djennane, S. (2003). 3D-Audio News Presentation Modeling. *Lecture Notes in Computer Science (LNCS)*, Vol. 4556, pp. 280-286, ISBN: 978-3-540-00855-2

Doermann, D. S.; Rivlin, E. & Rosenfeld, A. (1998). The function of documents. *Image Vision Computing*, Vol. 16, No. 11, 1 August 1998, pp. 799-814, ISSN: 0262-8856
Dormann, C. (2003). Affective experiences in the home: measuring emotions, *International Conference on Home Oriented Informatics and Telematics* (*HOIT2003)*, California, U.S.A., 6-8 April 2003
Edwards, A. D.; McCartney, H. & Fogarolo, F. (2006). Lambda: a multimodal approach to making mathematics accessible to blind students, *Proceedings of the 8th international ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '06)*, pp. 48-54, ISBN:1-59593-290-9, Portland, Oregon, USA, 23-25 October 2006, ACM, New York
Eglin, V. & Bres, S., (2003). Document page similarity based on layout visual saliency: Application to query by example and document classification, *Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR)*, Vol. 2, pp. 1208, ISBN: 0-7695-1960-1, Washington DC, 3–6 August 2003, IEEE Computer Society
Fellbaum, K. & Kouroupetroglou, G. (2008). Principles of Electronic Speech Processing with Applications for People with Disabilities. *Technology and Disability*, Vol. 20, No. 2, pp. 55-85, ISSN: 1055-4181
Fink, J.; Kobsa, A. & Nill, A. (1999). Adaptable and Adaptive Information Provision for All Users, Including Disabled and Elderly People. *New Review of Hypermedia and Multimedia*, Vol. 4, pp. 163-188, ISSN: 1361-4568
Fourli-Kartsouni, F. ; Slavakis, K. ; Kouroupetroglou, G. & Theodoridis S. (2007). A Bayesian Network Approach to Semantic Labelling of Text Formatting in XML Corpora of Documents. *Lecture Notes in Computer Science (LNCS),* Vol. 4556, pp. 299-308, ISBN: 978-3-540-73282-2
Francioni, J. M. & Smith, A. C. (2002). Computer science accessibility for students with visual disabilities, *Proceedings of the 33rd SIGCSE Technical Symposium on Computer Science Education (SIGCSE '02*), pp. 91-95, ISBN: 1-58113-473-8, Cincinnati, Kentucky, 27 February – 3 March 2002, ACM, New York
Freitas, D. & Kouroupetroglou, G. (2008). Speech Technologies for Blind and Low Vision Persons. *Technology and Disability*, Vol. 20, No. 2, pp. 135-156, ISSN: 1055-4181
Gorny, P. (2000). Typographic semantics of Webpages Accessible for Visual Impaired Users, Mapping Layout and Interaction Objects to an Auditory Interaction Space, *Proceedings of 7th International Conference on Computer Helping with Special Needs*, pp. 17-21, ISBN: 3-85403-145-9, Karlsruhe, Germany, 17-21 July 2000
Guillon B.; Monteiro J. L.; Checoury C.; Archambault D. & Burger D. (2004). Towards an Integrated Publishing Chain for Accessible Multimodal Documents. *Lecture Notes in Computer Science (LNCS)*, Vol. 3118, pp. 514-521, ISBN: 978-3-540-22334-4
Hall, R. H. & Hanna, P. (2004). The impact of web page text-background colour combinations on readability, retention, aesthetics and behavioural intention. *Behaviour & Information Technology*, Vol. 23, No. 3, May 2004, pp. 183-195, ISSN: 1362-3001
Harper, S. & Yesilada, Y. (2007). Web Authoring for Accessibility (WAfA). *Web Semantics: Science, Services and Agents on the World Wide Web*, Vol. 5, No. 3, pp. 175-179, ISSN:1570-8268
Harrington, S. J.; Naveda, J. F.; Jones, R. P.; Roetling, P.; & Thakkar, N. (2004). Aesthetic measures for automated document layout, *Proceedings of the 2004 ACM Symposium on Document Engineering (DocEng '04),* pp. 109-111, ISBN: 1-58113-938-1, Milwaukee, Wisconsin, USA, 28-30 October 2004, ACM, New York

Hill, A. & Scharff, L. V. (1997). Readability of screen displays with various foreground/background color combinations, font styles, and font types, *Proceedings of the Eleventh National Conference on Undergraduate Research (NCUR-97)*, Vol. 2, pp. 742-746, Austin, Texas U.S.A., 24-26 April 1997

Hjalmarsson, A., (2005). Adaptive Spoken Dialog Systems. In *GSLT, Speech Technology 1 Closing Seminar*

Holmberg, N. (2004). *Eye movement patterns and newspaper design factors. An experimental approach*, Master Thesis, Lund University Cognitive Science, Lund: LUCS, ISSN: 1101-8453, Sweden

Holmqvist, K. & Wartenberg, C. (2005). *The role of local design factors for newspaper reading behaviour – an eye-tracking respective*, Lund University Cognitive Studies, No. 127, Lund: LUCS, ISSN: 1101-8453, Sweden

Humaine Portal (2008), *http://emotion-research.net/*

IELA (2008). Initiative for Equitable Library Access, *http://www.lac-bac.gc.ca/iela/*

InkML, (2003). Ink Markup Language, Requirements for Ink Markup Language, *http://www.w3.org/TR/inkreqs/*

InkML, (2006). Ink Markup Language, *http://www.w3.org/2002/mmi/ink/*

ISO, (1989). Information processing -- Text and office systems -- Office Document Architecture (ODA) and interchange format, *http://www.iso.org/iso/iso_catalogue/catalogue_ics/catalogue_detail_ics.htm?csnumber=15926/*

ITU, (1993). ITU-T T.412, Open document architecture (ODA) and Interchange format – document structures, *Telecommunication (03/93) Standardization Sector of ITU Information Technology*

Kärkkäinen, L. & Laarni, J., (2002). Designing for small display screens, *Proceedings of the Second Nordic Conference on Human-Computer interaction,* Vol. 31, pp. 227-230, ISBN: 1-58113-616-1, Aarhus, Denmark, 19–23 October 2002, NordiCHI '02, ACM, New York

Kouroupetroglou, C.; Salampasis, M. & Manitsaris, A. (2007). Browsing shortcuts as a means to improve information seeking of blind people in the WWW. *Universal Access in the Information Society*, Vol. 6, No. 3, pp. 273-283, November 2007, ISSN: 1615-5289

Kramer, G., (Ed) (1994). *Auditory Display: Sonification, Audification, and Auditory Interfaces*, Addison-Wesley Longman Publishing Co., Inc., ISBN: 0201626039, Boston, MA, USA

Krishnamoorthy, M.; Nagy, G.; Seth, S. & Viswanathan, M. (1993). Syntactic segmentation and labeling of digitized pages from technical journals. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 15, No. 7, July 1993, pp. 737-747, ISSN: 0162-8828

Kupper, N. (1989). Recording of Visual Reading Activity: Research into Newspaper Reading Behaviour, *(Available as pdf from http://calendardesign.de/leseforschung/ /eyetrackstudy.pdf)*

Laarni, J. (2003). Effects of color, font type and font style on user preferences, *Adjunct Proceedings of HCI International 2003*, C. Stephanidis (Ed.), pp. 31-32, Heraklion, Greece, ISBN: 960-524-166-8, Crete University Press

Larson, K. (2007). The Technology of Text. *IEEE Spectrum*, Vol. 44, No. 5, May 2007, pp 26-31, ISSN: 0018-9235

Mao, S; Rosenfeld, A. & Kanungo, T. (2003). Document structure analysis algorithms: a literature survey. *Proceedings of SPIE Electronic Imaging*, Vol. 5010, pp. 197-207, January 2003, ISBN 0-8194-4810-9

Masanori, K.; Takeshi, M.; Eizen, K. & Kazuhito, U. (2007). Proposal of BrailleML as an XML for Japanese Braille: Conversion from ODF to BrailleML. *IEIC Technical Report,* Vol. 106, No. 485, pp. 31-36, ISSN: 0913-5685

MathPlayer (2008). MathPlayer, Design Science, *http://www.dessci.com/en/products/mathplayer/*

McTear, M. F. (2002). Spoken dialogue technology: enabling the conversational user interface. *ACM Computing Surveys (CSUR),* Vol. 34, No. 1, March 2002, pp. 90-169, ISSN: 0360-0300

Mills, C. B. & Weldon, L. J. (1987). Reading Text from computer screens. *ACM Computing Surveys (CSUR)*, Vol. 19, No. 4, December 1987, pp. 329–357, ISSN: 0360-0300

Mikhalenko, P. V. (2008), Multimodal interaction promises device integration, accessibility, and enhanced communication services, *http://articles.techrepublic.com.com/5100-10878_11-5090322.html*

Mynatt, E. D. (1994) Designing with auditory icons: how well do we identify auditory cues? *Conference Companion on Human Factors in Computing Systems,* pp. 269-270, ISBN: 0-89791-651-4, Boston, Massachusetts, USA, 24-28April 1994, C. Plaisant, Ed., CHI '94, ACM Press, New York

ODF (2008a). Open Document Format, *http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=office*

ODF (2008b). Open Document – Accessibility, *http://www.oasis-open.org/committees/ /tc_home.php?wg_abbrev=office-accessibility*

OASIS (2008). Organization for the Advancement of Structured Information Standards, *www.oasis-open.org/home/index.php*

Peels, A. J.; Janssen, N. J. & Nawijn, W. (1985). Document architecture and text formatting. *ACM Transactions on Information Systems (TOIS),* Vol. 3, No. 4, October 1985, pp. 347-369, ISSN: 1046-8188

Porat, T.; Liss, R. & Tractinsky, N. (2007). E-Stores Design: The Influence of E-Store Design and Product Type on Consumers' Emotions and Attitudes. *Lecture Notes in Computer Science (LNCS)*, Vol. 4553, pp. 712-721, ISBN: 978-3-540-73109-2

Power, C. D. (2008). Multi-Modal Exploration, PhD *thesis*, University of Western Ontario

Pitrelli, J. F.; Bakis, R.; Eide, E. M.; Fernandez, R.; Hamza, W. & Picheny, M. A. (2006). The IBM expressive Text-to-Speech synthesis system for American English. *IEEE Transactions on Audio, Speech and Language Processing*, Vol. 14, No. 4, July 2006, pp. 1099-1108, ISSN: 1558-7916

Raman, T. V. (1992). An Audio view of (LA)TEX Documents, *Proceedings of the Annual Meeting TUGboat*, Vol. 13, No. 3, pp. 65–70, ISBN: 0896-3207, Portland, Oregon, October 1992, TEX Users Group, USA

Richard, H. & Patrick, H. (2004). The Impact of Web Page Text-Background Colour Combinations on Readability, Retention, Aesthetics and Behavioural Intention. *Behaviour and Information Technology*, Vol. 23, No. 3, May-June 2004, pp.183-195, ISSN-0144-929X

Rosmaita, B. J. (2006). Accessibility first!: a new approach to web design, *Proceedings of the 37th SIGCSE Technical Symposium on Computer Science Education (SIGCSE '06)*, pp. 270-274, ISBN: 1-59593-259-3, Houston, Texas, USA, 3-5 March 2006, ACM, New York

RPH (2008). Radio for the Print Handicapped, *http://www.rph.org.au/*

Saari, T.; Turpeinen, M.; Laarni, J.; Ravaja N. & Kallinen, K. (2004). Emotionally Loaded Mobile Multimedia Messaging. *Lecture Notes in Computer Science (LNCS),* Vol. 3166, pp. 476-486, ISBN: 978-3-540-22947-6

Saito, S.; Takagi, H., & Asakawa, C. (2006). Transforming flash to XML for accessibility evaluations, *Proceedings of the 8th international ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '06)*, pp. 157-164, ISBN: 1-59593-290-9, Portland, Oregon, USA, 23-25 October 2006, ACM, New York

Sánchez, J. A.; Kirschning, I.; Palacio, J. C. & Ostróvskaya, Y. (2005). Towards mood-oriented interfaces for synchronous interaction, *Proceedings of the 2005 Latin American Conference on Human-Computer interaction*, Vol. 124, pp. 1-7, ISBN:1-59593-224-0, Cuernavaca, Mexico, 23-26 October 2005, CLIHC '05, ACM, New York

Sánchez, J. A.; Hernández, N. P.; Penagos J. C. & Ostróvskaya, Y. (2006). Conveying mood and emotion in instant messaging by using a two-dimensional model for affective states, *Proceedings of the Symposium on Human Factors in Computer Systems IHC 2006*, pp. 66-72, ISBN: 1-59593-432-4, Brazil, 2006, ACM, New York

Schamber, L., (1996). What is document? Rethinking the concept in uneasy times. *Journal of the American Society for Information Science*, Special issue: Electronic Publishing, Vol. 47, No. 9, pp. 669–671, ISSN: 0002-8231

Schröder, M. (2006). Expressing degree of activation in synthetic speech. *IEEE Transactions on Audio, Speech and Language Processing*, Vol. 14, No. 4, July 2006, pp. 1128-1136, ISSN: 1558-7916

Soiffer, N., (2005), MathPlayer: Web-based Math Accessibility, *Proceedings of the 7th international ACM SIGACCESS Conference on Computers and Accessibility (ASSETS '05)*, pp. 204-205, ISBN: 1-59593-159-7, Baltimore, MD, USA, 9-12 October 2005, ACM, New York

Spiliotopoulos, D.; Xydas, G.; Kouroupetroglou, G. & Argyropoulos, V. (2005a). Experimentation on Spoken Format of Tables in Auditory User Interfaces, *Proceedings of the 11th International Conference on Human-Computer Interaction (HCII2005)*, pp. 361-370, ISBN: 0-8058-5807-5, Las Vegas, Nevada, USA, 22-27 July 2005, Lawrence Erlbaum Associates, Inc

Spiliotopoulos, D.; Xydas, G. & Kouroupetroglou, G. (2005b). Diction Based Prosody Modeling in Table-to-Speech Synthesis. *Lecture Notes in Artificial Inteligence (LNAI)* Vol. 3658, September 2005, pp. 294–301, ISBN: 978-3-540-28789-6

Stephanidis, C. (2001). *User Interfaces for All. Concepts, Methods, and Tools*. Lawrence Erlbaum Associates, Inc., Publishers, ISBN: 0-8058-2967-9, USA

SVG (2000), Scalar Vector Graphics, Accessibility Features of SVG, *http://www.w3.org/TR/SVG-access/*

SVG (2008), Scalar Vector Graphics, XML Graphics for the Web, *http://www.w3.org/Graphics/SVG/*

Tsonos, D.; Xydas, G. & Kouroupetroglou G. (2007a). Auditory Accessibility of Metadata in Books: A Design for All Approach. *Lecture Notes in Computer Science (LNCS)*, Vol. 4556, pp. 436 - 445, ISBN: 978-3-540-73282-2

Tsonos, D.; Xydas, G. & Kouroupetroglou, G. (2007b). A Methodology for Reader's Emotional State Extraction to Augment Expressions in Speech Synthesis, *Proceedings of 19th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2007)*, Vol. 2, pp. 218-225, ISBN: 0-7695-3015-X, Patras, Greece, 29-31 October 2007

Tsonos, D. & Kouroupetroglou, G. (2008). Accessibility of Board and Presentations in the Classroom: a Design-for-All Approach, *IASTED International Conference on Assistive Technologies (AT 2008)*, pp. 13-18, ISBN: 978-0-88986-739-0, Baltimore, Maryland, USA, 16-18 April 2008

Tsonos, D.; Ikospentaki, K. & Kouroupetrglou, G. (2008). Towards Modeling of Readers' Emotional State Response for the Automated Annotation of Documents, *IEEE World Congress on Computational Intelligence (WCCI 2008)*, pp. 3252 - 3259, ISSN: 978-1-4244-1821-3, Hong Kong, 1–6 June 2008

Tsujimoto, S. & Asada, H. (1990). Understanding multi-articled document, *Proceedings of 10th International Conference on Pattern Recognition*, Vol. 1, pp. 551-556, ISBN: 0-8186- 2062-5, Atlantic City, NJ, USA, 16-21 June 1990

Turunen, M.; Hakulinen, J.; Räihä, K.; Salonen, E.; Kainulainen, A. & Prusi, P. (2005). An architecture and applications for speech-based accessibility systems. *IBM Systems Journal*, Vol. 44, No. 3, August 2005, pp. 485-504, ISSN: 0018-8670

W3C (2008a). XML Document Navigation Language, *http://www.w3.org/TR/xdnl/*

W3C (2008b). Multimodal Interaction Activity, *http://www.w3.org/2002/mmi/*

W3C (2008c). World Wide Web Consortium, *www.w3.org*

W3C (2008d). Web Content Accessibility Guidelines 1.0 (WCAG 1.0), *http://www.w3.org/TR/WCAG10/*

W3C (2008e). Web Accessibility Initiative (WAI), *http://www.w3.org/WAI/*

W3C (2008f). W3C Math Home, *http://www.w3.org/Math/*

W3C (2008g). Mathematical Markup Language (MathML) Version 2.0 (Second Edition), Introduction, *http://www.w3.org/TR/2003/REC-MathML2-20031021/chapter1.html*

Wartenberg, C. & Holmqvist, K. (2005). *Daily Newspaper Layout - Designers' Predictions of Readers' Visual Behaviour - A Case Study*, Lund University Cognitive Studies, No. 126, Lund: LUCS, ISSN: 1101-8453, Sweden

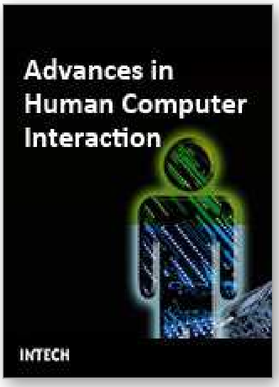
Xydas, G. & Kouroupetroglou, G. (2001a). Text-to-Speech Scripting Interface for Appropriate Vocalisation of e-Texts. *Proceedings of EUROSPEECH 2001*, pp. 2247–2250, ISBN: 87-90834-09-7, Aalborg, Denmark, 3-7 September 2001, International Speech Communication Association

Xydas, G. & Kouroupetroglou, G. (2001b). Augmented Auditory Representation of e-Texts for Text-to-Speech Systems. *Lecture Notes in Artificial Intelligence (LNAI)*, Vol. 2166, pp. 134-141, ISBN: 978-3-540-42557-1

Xydas, G.; Spiliotopoulos, D. & Kouroupetroglou, G. (2003). Modelling Emphatic Events from Non-Speech Aware Documents in Speech Based User Interfaces, *Proceedings of HCI International 2003 - The 10th International Conference on Human-Computer Interaction*, pp. 806–810, ISBN: 0805849319, Crete, Greece, 22-27 June 2003, C. Stephanidis and J. Jacko, eds,. Lawrence Erlbaum Associates, Inc., Mahwah, NJ

Xydas, G.; Argyropoulos, V.; Karakosta, T. & Kouroupetroglou, G. (2005). An Experimental Approach in Recognizing Synthesized Auditory Components in a Non-Visual Interaction with Documents, *Proceedings of the 11th Int. Conference on Human-Computer Interaction*, Vol. 3, pp. 411-420, ISBN: 0-8058-5807-5, Las Vegas, Nevada, USA, 22-27 July 2005, Lawrence Erlbaum Associates, Inc

Yamashita, A.; Amano, T.; Takahashi, I. & Toyokawa, K. (1991). A model based layout understanding method for the document recognition system, *Proceedings of International Conference on Document Analysis and Recognition*, pp. 130-138, Saint Malo, France, September 1991



Advances in Human Computer Interaction

Edited by Shane Pinder

ISBN 978-953-7619-15-2

Hard cover, 600 pages

Publisher InTech

Published online 01, October, 2008

Published in print edition October, 2008

In these 34 chapters, we survey the broad disciplines that loosely inhabit the study and practice of human-computer interaction. Our authors are passionate advocates of innovative applications, novel approaches, and modern advances in this exciting and developing field. It is our wish that the reader consider not only what our authors have written and the experimentation they have described, but also the examples they have set.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Georgios Kouroupetroglou and Dimitrios Tsonos (2008). Multimodal Accessibility of Documents, Advances in Human Computer Interaction, Shane Pinder (Ed.), ISBN: 978-953-7619-15-2, InTech, Available from: http://www.intechopen.com/books/advances_in_human_computer_interaction/multimodal_accessibility_of_documents

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821