

“Braillekrant” and “DiGiKrant”: a Daily Newspaper for Visually Disabled Readers

Bert Paepen, Jan Engelen

Katholieke Universiteit Leuven—Research Group on Document Architectures
Kasteelpark Arenberg 10—B-3001 Leuven
bert.paepen@esat.kuleuven.ac.be

Abstract

At present Belgium is one of the rare countries that have a daily newspaper accessible to visually impaired readers. Both an electronic version (DiGiKrant) and a Braille paper version (BrailleKrant) are published on a day-to-day basis. This paper shows what challenges exist for visually impaired readers to read a paper and how these challenges can be overcome. It further explains how the accessible newspapers are being created. In the publication process the XML standard is extensively used, not only as an exchange format between content providers and the publishers, but also as a format enabling a usable access to the electronic newspaper by the visually impaired reader. The paper shows which information formats are used and how the process of conversion between them works.

1 Introduction

The main goal of the “Braillekrant” and “DiGiKrant” projects is to offer an accessible daily newspaper to visually disabled readers in Flanders. It is very important that disabled readers can read the newspaper in an independent way, so without the help of a visually enabled reader. To achieve this goal the non-profit organization “Braillekrant VZW” [1] has been founded. In cooperation with the publishing company VUM (Vlaamse Uitgevers Maatschappij), a “blind-friendly” version of most VUM newspapers is being published: De Standaard, Het Nieuwsblad and Het Volk.

The software for creating the accessible newspapers from the original source as provided by the publisher has been contrived and implemented at the Research Group on Document Architectures – DocArch [2] – of the Katholieke Universiteit Leuven [3] and has its foundation in several research projects. This paper will first introduce the project and explain how a visually disabled person can read a newspaper. Then the paper explains how the information flow is going in the software that transforms the original source files to the format readable by the end user. The paper concludes with an overview of problems—encountered during the project—that can happen between the authoring of an article and the publication in an accessible form.

2 BrailleKrant V.Z.W.

At the initiative of two K.U.Leuven students, who experienced a great lack of written information for visually impaired readers, the organisation BrailleKrant (DBK) was founded in 1990. After the idea had emerged to publish a daily printed newspaper in Braille, several projects were initiated to make this idea technically feasible. About two years later the first Braille copy of the quality newspaper De Standaard was published in cooperation with the publishing group VUM [4]. In 1998 the popular newspaper Het Nieuwsblad was added to the Braille publication list.

Since 1997 an additional publication channel was created, called the “DiGiKrant”: a digital version of the newspaper, sent to subscriber using a floppy disk or, in a later phase, e-mail. The main advantage of the digital paper is that it contains *all* the news articles from the original newspaper, whereas the Braille paper only contains a *selection* of about 20 key articles. The reason for this limited selection lies in the Braille format that needs quite thick paper and that takes up a lot of space on the paper. (Figure 1) shows one copy of the Braille paper (Het Nieuwsblad 2003-06-18) containing only 20 articles, but using 32 double-sided sheets of thick paper. The same information printed in Courier font on regular paper uses only 1/4 of the amount of sheets. Note that in (Figure 1) the Braille copy is double sided and the regular printed copy is single sided.

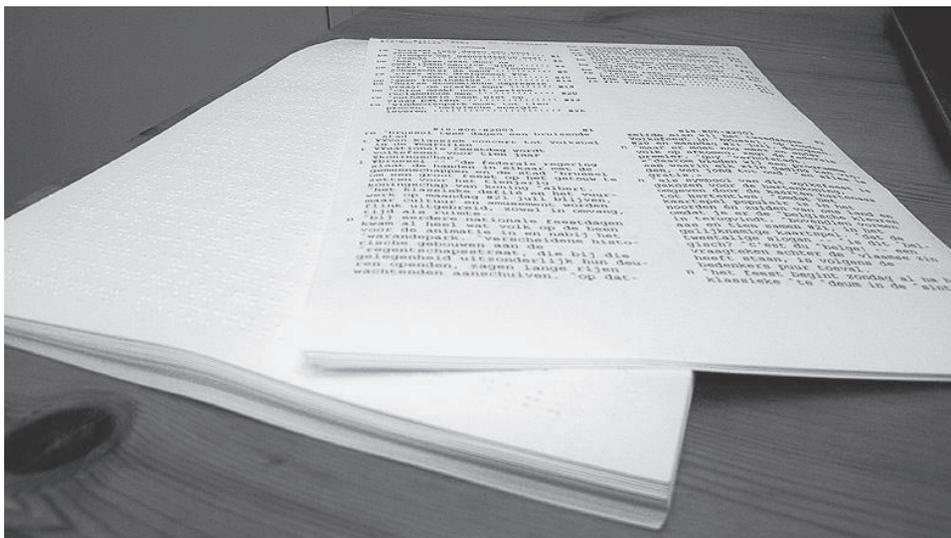


Figure 1: Newspaper size in Braille versus regular print

3 Concept of the Accessible Newspaper

The goal of this paper is not to explain in detail the problems experienced by visually impaired readers, but rather to show how information structures can be used to make an accessible electronic newspaper.

3.1 Solutions for Accessible Reading

A lot of visual impairments can make it difficult to process written information – let’s keep the newspaper as an example. A reader with partial sight might not be able to read the smaller letters, but he could use a magnifying glass to be able to read the entire newspaper. Other reading impairments include blindness, dyslexia, colour-blindness and motoric dysfunctions. Where magnifying glasses are quite primitive tools, much more sophisticated solutions exist that help impaired readers surpass their limitations.

Screen magnifiers can be compared to the traditional magnifying glasses; they can be used for enlarging letters shown on the screen. Text-to-speech software transforms written text directly to a computer-generated voice, so that you can listen to the written text as if someone reads it for you. Finally a screen reader is a device that makes the Braille format—with all its limitations like the massive amount of paper needed—more flexible. When connected to a computer the device can display one line of text on one line of Braille characters. The reader can then, line by line, read the text shown on the computer screen. These three examples are far from exhaustive but they summarise the main categories of solutions for reading the DiGiKrant [5].

3.2 Reading Challenges

Even with all this equipment at hand, handling information remains a cumbersome task. These applications and devices typically offer information in very small chunks, for example line-by-line (screen readers) or on-screen area-by-area (magnifiers). As a consequence, visually impaired readers are forced to handle electronic information in parts; “seeing” the big picture at once mostly is impossible. For web sites this is not a big problem, since they tend to chop up their information in smaller parts using various ways of navigation. For large texts such as books, newspapers and journals, a much bigger problem exists. Large amounts of text are hardly ever presented in a structured way. They can contain chapters and paragraphs, like this paper does, but that does not mean that their structure is *offered* to the reader in a structured way.

People reading a printed newspaper typically have a quick look at the large headlines, allowing them to browse through the entire paper in a couple of minutes. In this example, formatting helps people cope with large amounts of information. Not only the amount of text can be an obstacle for visually impaired persons, but also the way this text is presented. For some types of reading impairment, like partial sightedness or colour-blindness, formatting can make reading easier, for example by using specific colour combinations and large titles. However for blind readers formatting cannot help in any way.

If a blind reader receives a digital copy of the same newspaper without special treatment, quick browsing is impossible. He or she would have to read the newspaper line-by-line, not being able to see the difference between big and small headlines and not being able to skip uninteresting parts. [6]

3.3 XML can Make Information More Accessible

XML (Extensible Markup Language) [7] brings an important principle to the accessibility world: the separation of information from its presentation form. If you maintain this separation, it becomes easy to create different presentation forms for different kinds of impairment. If personalised presentation of information is often seen as a gimmick, for a lot of impaired readers it is a necessity.

Making information more accessible thus means restructuring both the information and the way it is presented to the reader. In order to make large amounts of information accessible to visually impaired persons, they need to be offered in a structured way, so that disadvantages of line-by-line viewing are diminished. Further, navigation and search must be enabled in a user-friendly way. Information must be presented in manageable chunks and the way it is presented must be very flexible and user-adaptable [6]. In other words: the *limitations* imposed by accessibility software and hardware have to be transferred to *requirements* for the information structure and the way this information is presented so that they can become *enablers* for impaired readers.

For DiGiKrant the DAISY format is used, an XML based standard format for digital talking books. Several types of books can be stored using this format: audio books, textbooks and several audio-text combinations with links between the audio and the text (for details see [9]). Whatever form is used, the DAISY format enables readers to view the hierarchical structure of the document and to quickly navigate through the document, for example by skimming to the next paragraph. DAISY formatted books are independent from any reading application.

Next to the accessible presentation of information XML serves another goal in this project: the information exchange between the content providers' databases and the DBK conversion software. That way the conversion software operates independently of the providers' database format and possible changes to it.

3.4 Requirements of an Accessible Newspaper

When can a newspaper be called “accessible” for an impaired reader? A newspaper narrated to a blind person by a friend is accessible. A narrated newspaper stored on an audio cassette is also accessible. But in both cases the usefulness is quite low. In the former case the reader depends on the goodwill of the friend, while in the latter case the reader depends on the order and speed in which the newspaper was recorded on the audio disc.

For making a newspaper really accessible and usable the reading experience should be an individual one and one where the reader is in charge all the time, deciding what to read and what to skip. In other words the reading experience should be almost equal to the reading of a traditional newspaper. This also means that impaired readers should – as far as technically feasible – get exactly the same information as offered in a regular paper.

The blind-friendly newspaper is being published on a daily basis in two forms: Braille and electronic. The Braille version is printed on paper using a dedicated embosser and is sent during the night to all subscribers; the electronic version is sent to its subscribers by e-mail or on floppy disk. For both versions this means that the visually disabled subscribers receive their newspaper at the same moment as the subscribers of the traditional newspaper – sometimes even quicker. This is a very important requirement in the *access for all* discussion.

3.5 DiGiKrant Reading Software

In its current status the DiGiKrant uses the software application Kurzweil 1000 from Kurzweil Educational Systems to make the electronic newspaper usable for impaired readers. Kurzweil 1000 turns text into a variety of life-like voices. Users can bookmark, make notes, skip or summarize text to aid mastery of concepts and details. They can modify reading speed, pitch, and emphasis to suit specific preferences and purposes. [8] Through the use of style sheets various reading impairments can be addressed. Next to the text-to-speech mode, users can also use their own software or hardware to make the information accessible.

When a newspaper is opened in the reading application the general structure of the newspaper is shown in a tree-like table of contents. The first level is used for the sections of the paper, like “front page”, “economy”, “sports”, etc. The next levels are used for subsections and for articles. One simple feature increased the usability of the electronic paper: adding the number of articles contained in a section between brackets behind the section name. That way the reader has a quicker overview when only the sections are shown (see left part of Figure 2). The right part of (Figure 2) shows the articles contained in a section.

A second feature made the computer-narrated paper more usable: adding the text [article end] to the end of each article. That way the transition from one article to the other becomes clearer.

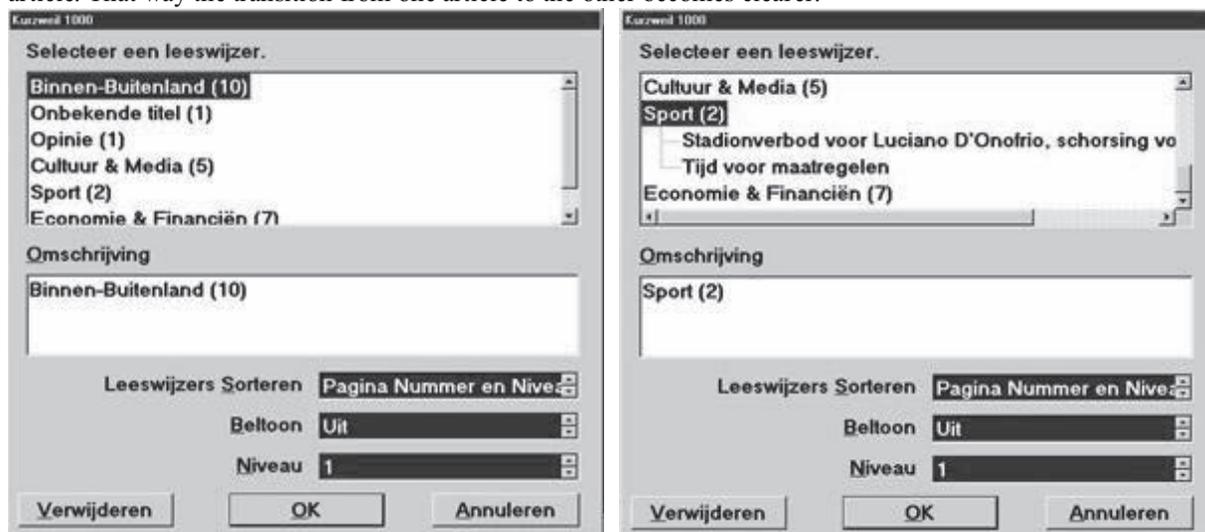


Figure 2: DiGiKrant table of contents view in Kurzweil 1000

4 Information Flow from Author to Reader

While all newspapers offered by BrailleKrant VZW (DBK) originate from the publisher VUM, the publication of the accessible newspapers is done by DBK employees. Before the accessible newspapers can be published in a readable form, a lot of steps have to be taken starting from the original articles in the VUM databases. In order to make this complicated process as simple as possible a dedicated conversion wizard has been developed. This application collects all the necessary data, takes care of all the conversion steps and publishes the newspapers to all subscribers in the correct format.

4.1 Formats and Data Flow

To get a view on the complex publication process (Figure 3) illustrates the variety of data formats that are handled by the conversion wizard and the data flows managed by the conversion wizard.

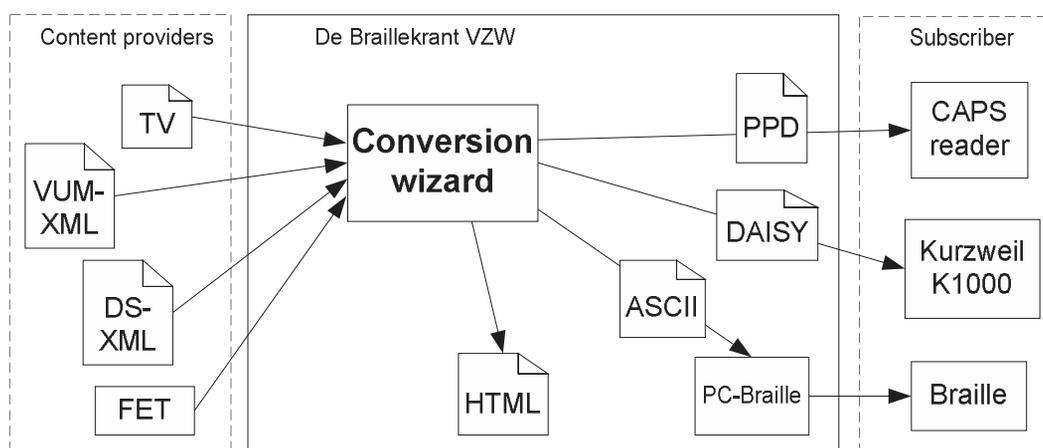


Figure 3: Data flow for DiGiKrant and BrailleKrant publication

The following input formats are being used:

- TV: ASCII-based format for TV guide information
- VUM-XML: format exported by VUM prepress database; used for DiGiKrant publication
- DS-XML: format exported by VUM web database; used for BrailleKrant publication
- FET: format used for financial information like stock exchange tables

The conversion wizard can generate the following output formats:

- PPD: format for DiGiKrant subscribers using the DOS reader CAPS (predecessor of K1000)
- DAISY: format for DiGiKrant subscribers using the Kurzweil 1000 reader
- ASCII: format for publication to Braille embosser; for this publication the intermediate application PC-Braille is used as the Braille format requires a very specific page layout.
- HTML: version of the newspaper that is publishable to a website

The conversion wizard is designed in such a way that all input formats can be converted to all possible output formats. Regardless the format the publication can be on a floppy disk sent to the user overnight or using e-mail. In the former case the subscriber receives the paper in the morning; in the latter case he or she already receives tomorrow’s paper shortly after midnight.

One of the main problems the conversion wizard is faced with is the difference between the basic information structure of the input and the output files. While the input formats VUM-XML and DS-XML are single article based, the output formats are newspaper-based. This means that the conversion software has to entirely reconstruct the structure of the original newspaper based on individual article files. Another issue is the order of news articles within one section (and equally the order of sections within the newspaper). On a printed newspaper it is clear which articles are more important than others, because it shows in the size of their title, but this information is not contained at all in the source XML files of the articles. Therefore articles in DiGiKrant cannot be ordered within each section according to their importance.

4.2 BrailleKrant Publication process

Being an electronic newspaper, the creation and publication of the DiGiKrant happens in an almost completely automated process, depicted in (Figure 3). For the Braille version (BrailleKrant) however, more manual intervention is necessary.

As the amount of information contained on one Braille page is limited (each Braille embossed character takes a lot of space in comparison with a printed character), the Braille newspaper only contains a selection of about 20 articles taken from the traditional newspaper. This selection is the first publication step, which is done by the editors of DBK using the Backoffice tool (see Figure 4). In practice some effort is reused from the editorial office of the online VUM newspapers. In order to allow priority listings of important and less important news articles on their websites, these editors define a priority for each article. These priority numbers are reused for determining which articles should be selected for the BrailleKrant and which not. Only the selected articles are then exported to the DS-XML format on the FTP server (see Figure 4). This priority listing is the only reason why a different format is used as the source for DiGiKrant (VUM-XML) versus BrailleKrant (DS-XML).

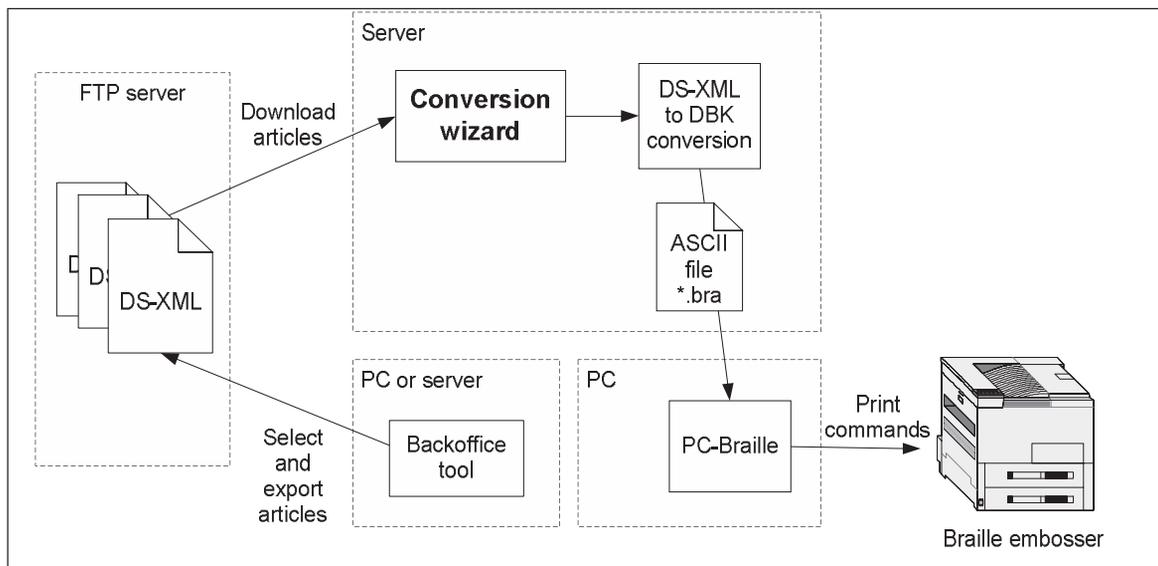


Figure 4: BrailleKrant publication process

After this export has been made, the second publication step takes place. The conversion wizard collects the DS-XML files, reconstructs a newspaper out of them and converts this to an ASCII format that can be used in the

application PC-Braille. This application is dedicated to the kind of Braille embosser used and transforms the ASCII newspaper into print commands for the embosser, a machine that is able to print Braille on paper. This is the final publication step. In PC-Braille the final edition is done: the text is spread over a number of Braille pages, page numbers are added and finally a table of contents is added with references to the correct pages. The Braille papers (and equally the electronic papers on floppy disk) are then sent to the subscribers by surface mail at the same time as the ordinary newspapers.

5 Conclusions and Lessons Learned

One of the main problems encountered during the development of the conversion software was the lack of metadata. One problem was already mentioned before—the lack of priority indications for news articles—but even a more fundamental problem is the lack of title information for a lot of articles. A test revealed that for the analysed newspapers between 5% and 12% of all articles delivered in the VUM-XML format did not have a title. Because titles are very important for assessing the article's importance and relevance to the reader, the decision was made to automatically generate titles for those articles, using the first limited number of characters from the article text.

The base problem related to this is the fact that the source files originate from a database that is completely targeted toward the paper publication of news. Aspects that are important for the electronic publication or conversion are therefore not handled properly by this database. This example shows that for this kind of publication it is very important to adhere to the single source principle as much as possible: only one database should be used for print, online publication and derived products such as BrailleKrant and DiGiKrant. That way, if corrections are applied for the benefit of the one publication (like manually adding missing titles) they can be easily reused for other purposes.

References

- [1] Braillekrant VZW. <http://www.braillekrant.be>
- [2] Bert Paepen et.al. <http://www.docarch.be>
- [3] <http://www.kuleuven.be>
- [4] Braillekrant VZW. De BrailleKrant v.z.w.—over ons, <http://www.braillekrant.be/overons.html> , 2005-04-28
- [5] Braillekrant VZW. De BrailleKrant—Kranten voor personen met een leeshandicap, <http://www.braillekrant.be/kranten.php> , 2005-04-28
- [6] Bert Paepen and Jan Engelen (2002). Using XML as a Reading Enabler for Visually Impaired Persons. In K. Miesenberger, J. Klaus, W. Zagler (Eds), ICCHP 2002, LNCS 2398, pp. 382-389, 2002. Springer-Verlag Berlin Heidelberg.
- [7] Liam Quin. Extensible Markup Language (XML), <http://www.w3.org/XML/>, 2005-04-29
- [8] Kurzweil Educational Systems. Kurzweil 1000 Overview—Reading Software for the blind and visually impaired, http://www.kurzweiledu.com/products_k1000.asp, 2005-04-29
- [9] Daisy Consortium. Technology Overview—What is a DTB?, http://www.daisy.org/about_us/dtbooks.asp, 2005-04-29