Digitale Archivering in Vlaamse Instellingen en Diensten

DAVID

Digital archiving of the electoral register

Filip Boudrez

Sofie Van den Eynde

FACULTEIT RECHTSGELEERDHEID

INTERDISCIPLINAIR CENTRUM VOOR RECHT EN INFORMATICA

TIENSESTRAAT 41

B-3000 LEUVEN

Stadsarchief Stad Antwerpen

Version 1.0

Legal deposit library D/2001/9.213/2

Antwerpen-Leuven, January 2001

DAVID is a research project financed by the Foundation for Scientific Research
in the scope of the Max wildiers Fund

TABLE OF CONTENTS

I. INTRODUCTION

It was decided that the electoral register should be the subject of the first DAVID[1] project case-study. In Belgium legislation governing elections requires that an electoral register be drawn up for each election and that that register contain the names of all those entitled to vote. This list serves a practical purpose when it comes to organizing the election operations on election day. Before the elections, the electoral register serves as a means for the public to check the identity of the electors who will be called upon to vote in the election and to give them the opportunity to express any objections to entries in the register.

The choice of the electoral register as the subject of the first DAVID project case-study can be justified on three counts: the assumption that the digital archiving of this document raises no problems of a juridical nature, the statutory obligation to permanently preserve the electoral register and the immediate need of the archival service of the City of Antwerp to archive the digital electoral register of the last elections (on October 8th 2000). As in most Flemish municipalities, there is both a paper and a digital electoral register in the City of Antwerp. When extended to its full length, the version on paper runs to approximately three meters.

This report endeavours to provide an answer to the following questions[2]. Is it lawful to produce the electoral register digitally and to archive it? Can the paper electoral register be destroyed and is it enough to archive just the digital electoral register? What requirements must the archiving of the digital electoral register meet? In what file format and on what storage medium will it be archived? How can the authenticity and integrity of this digital file be guaranteed? Can the archiving of the digital electoral register serve as a model for other digital files?

II. DIGITAL ARCHIVING OF THE ELECTORAL REGISTER: LAWFUL OR UNLAWFUL?

Initially the electoral register was drawn up manually each year from 'the municipal register', i.e. the list of all those living in the municipality. Those entitled to vote were selected from it and the result was the paper electoral register. With the advent of computerization, in the 1970s more and more electoral registers were produced digitally, drawing on the most recent data from the municipal

---

1 Digitale Archivering in Vlaamse Instellingen en Diensten (DAVID, or Digital Archiving in Flemish Institutions and Services) is financed by the Fund for Scientific Research as part of the Max Wildiers Fund. The DAVID project's two partners are the Interdisciplinair Centrum voor Recht en Informatica (ICRI, or Interdisciplinary Centre for Law and Information Technology) and the Antwerp City Archive.

2 This text is a summary of the full report produced in Dutch: F. BOUDREZ and S. VAN DEN EYNDE, *De digitale archivering van het kiezersregister* (The Digital Archiving of the Electoral Register), Antwerp-Leuven, 2001, 49 p. The complete report and the DTDs used can be downloaded from Http://www.antwerpen.be/david

register. The electoral legislators in Belgium were anxious to capitalize on the advantages of computer applications (fewer mistakes, speed of producing the electoral register) and to give them a legal footing, and in 1991 a ruling was introduced whereby only one electoral register can be drawn up for each election. Consequently, the concept of digitalization is embodied in the legislation governing elections. Hence the assumption in this report that there would be no obstacles to digital archiving.

Research into requirements laid down by the legislator with regard to the form the electoral register should take, shows that digitalization is indeed possible from a legal viewpoint. In fact, the requirements with regard to form only relate to the (structure of the) contents of the electoral register. For the electoral register there are no typical paper requirements with regard to form that make it necessary to preserve the paper document itself. When it comes to compiling the electoral register, the law stipulates two formal requirements, namely a number of obligatory entries such as the elector's family name and first name, and a specific structure according to which the electoral register has to be drawn up. The register must be compiled on a consecutive numbering basis and per district of the municipality or in alphabetical order of the electors.

The requirement that obligatory entries must appear in a document consisting of fields, poses no problem for digital archiving, provided it can be guaranteed that those entries also appear in the digital document. This formal requirement applies only to the content of the digital document.

The formal requirement of consecutive numbering relates to the elector's number; every person eligible to vote is assigned a number. So the electoral register can be used to check the number of voters that have to present themselves for the council elections. (Belgium is one of the few countries where voting is compulsory.) Likewise, consecutive numbering is a requirement that relates only to the contents of the document. So for digital archiving this requirement poses no problems either, provided that when drawing it up the correct number is allocated to the right elector and the link between the two data fields can be archived (just as the other obligatory entries have to be archived in such a way that their link remains with a specific elector).

The last requirement relating to form is the alphabetical or geographical order or sequence of the register. The sequence of the contents of a document is a typical paper requirement regarding form, and one that serves no purpose in a digital environment. In a digital environment the sequence of data can, after all, be altered very easily by means of a whole range of search and selection functions. Indeed, in a digital context the content of a document is no longer linked to the storage medium as it is in the paper world. Moreover, the order of the 'old-style' electoral register is not necessarily the order in which the data appear on the digital medium. Those data are certainly not in logical order, but the computer decides for itself where there is still space available to store the digital data and which arrangement is most economical. The sequence of data is, however, a typical paper requirement governing form, and it can be imitated in a digital environment. So from a legal point of view, there is no problem. This legal requirement of form is, however, best applied by deciding that it must be possible to put the document in a certain sequence if the electoral register is to be produced digitally.

III. THE INTEGRITY QUESTION

The electoral register, like all digital archival records, raises the question of integrity. Most notable about the paper electoral register are the special integrity measures relating to the use of the paper

medium, taken when compiling the electoral register. This we see if we consult the paper documents at Antwerp City Archive. Apart from the integrity-preserving character of paper itself as a medium, the way that paper is used can also contribute to a document's integrity. In other words, the paper registers are drawn up in such a way that there are no white areas, which means that nothing can be added to the electoral lists, or rather, it guarantees that any additions will be noticed immediately, because they appear in the margin, for example, or outside the normal printed page format. Secondly, all the pages of the paper electoral registers are numbered consecutively. This guarantees that any pages slipped in between are instantly noticed as this interferes with the consecutive numbering.

These two safeguards cannot be carried over into a digital environment in the same way because they lose their intended effect. It is not enough to ensure that the absence of white areas and the consecutive page numbers are part of the digital document. Those two elements would, after all, only be digital information, which is added to all the digital information that makes up the electoral register. If the digital electoral register is visualized on a screen, that visual reproduction will show the absence of white areas or the presence of page numbers. But as an entity made up of bits and bytes, the electoral register could still be altered at any time without anyone noticing, irrespective of whether the visual reproduction shows the absence of white areas or the page numbers.

With digital archiving, an alternative has to be found to the administrative measures concerning integrity. This report makes out a case for the technique of the digital signature as a solution to the integrity problem in a digital context.

IV. ACCESS TO THE ELECTORAL REGISTER

The electoral register contains a whole host of personal details. So basically it is a document to which access has to be restricted. The information contained in the electoral register may only be used for election purposes and may only be made public for those purposes to the people explicitly named in law, such as political parties. Private individuals and commercial enterprises (who would of course be very interested in these personal details for direct marketing purposes) would have great difficulty proving an electoral interest, especially once the elections are over and the electoral registers have been deposited in the archive, for even outside the period between the date of the compilation of the electoral list and the date of the election, the electoral registers may only be used for electoral purposes.

Yet one can conceive of situations where third parties must be allowed to have sight of the archived electoral register. In 1997 the Committee for the Safeguard of Personal Privacy decided that in some cases a special interest can take precedence over the safeguard of privacy; the interest, for example, of genealogists carrying out scientific research. So provision must be made for such people to consult the archived digital electoral register.

V. DESTRUCTION OF THE PAPER ELECTORAL REGISTER

One of the advantages of digital archiving for the archival services is the space-saving factor. This is of course only true if the paper versions of the documents are destroyed. Yet there is no unequivocal answer to the question of whether it is permissible to destroy the electoral register.

The problem is determining which is the original and approved digital version, and this is well nigh impossible to do. As it is not feasible to hand over the whole electoral register to the competent authorities for approval (for Antwerp alone we are talking about 316,926 electors), the number of voters and the number of polling stations is submitted for each district. The mayor and aldermen then ratify the electoral register in a single article. So it is not totally clear which particulars the authorities have approved.

The report argues that it is nevertheless right and proper to destroy the paper electoral register. If a digital electoral register is offered to the archival institution, then it can assume that that register is the original. It is after all the Administration that is responsible for the contents of the electoral register. Moreover, the approval procedure in the former paper environment was the same, so in a way there was no original then either.

VI. THE VALUE OF ARCHIVING THE ELECTORAL REGISTER

The report argues that the preservation of electoral registers in a digital context is totally unnecessary. First of all, the electoral register is a partial reproduction of the population registers. So in terms of content the electoral register is of no more import than the population registers. The arguments previously put forward by the Public Record Office in favour of preserving electoral registers, are no longer applicable either. The views outlined in this report will be put to the Public Record Office. The electoral register could then perhaps be removed from the preservation and destruction lists.

VII. DIGITAL ARCHIVING OF THE ELECTORAL REGISTER

The electoral register must be permanently preserved. The archiving of the digital electoral register must satisfy a number of requirements. The data file must be durable, usable, reliable and consultable and the legal requirements must be observed. In outlining an archiving strategy these requirements influence first and foremost the choice of file format and the storage medium.

A. THE DATA FILE FORMAT

When in future we want to consult a digital file, then the file must be compatible with the platforms and software applications being used at that particular time. To achieve this, two paths are open to us. Firstly, we can store the data file in the file format of the application in which it is made. When the required hardware and software becomes obsolete, the file has to be converted or even migrated. Such a solution requires meticulous file management, which is costly in the long term. Alternatively, the file can be stored in a format that is not application-linked and that will be compatible with the different platforms. The advantage of such a format is that it does not have to be moved from one place to another or converted when hardware and/or software becomes obsolete or is replaced. For the archiving of the digital electoral register, we advocate this second path. By consequence, the choice of file formats is limited to the file formats that are prescribed as standard.

Many so-called 'standards' are the product of one particular manufacturer and only have the status of *de facto* standards (e.g. Adobe's PDF, Microsoft's MS Office). The future usability of these formats is entirely in the hands of their manufacturer, which means that these formats are not suitable for long-term archiving. Suitable standards are those approved by an official or international standardization body. These standards have the advantage that they are supported by an organization that ensures that future developments are compatible with these standard formats. Both the *International Standard Organisation* (ISO) and the *World Wide Web Consortium* (W3C) are active in the IT field.

We decided upon XML for the archiving of the digital electoral register. XML (*eXtensible Mark-up Language*[3]) is a semantic descriptive mark-up metalanguage that is platform and vendor-independent. XML is a subset of SGML (*Standard Generalized Mark-Up Language: ISO 8879*) and is also able to automatically define tags, structure data fields and validate the structure. The principle of a semantic mark-up language is that content and layout are separate. Layout data are as a rule registered with their own application source language and usually pose problems when it comes to exchanging information between two different platforms. In the XML file only the content of a file is fixed. The content of a data field is placed between tags; these mark the beginning and the end of the field and tell us more about the meaning of the data (e.g. <date_of_birth>14041974</date_of_birth>). Thus semantics and data are archived together, which means people can understand XML files with the help of viewers.

[3] Documentation on XML can be found on: http://www.oasis-open.org/cover. We consulted: LANDER R., *XML: The new markup wave*; J. BOSAK, *XML, Java, and the future of the Web*; W. FIERZ et al, *The SGML standardization framework and the introduction of XML*; MACHERIUS I., *Expert's revolution. XML: a professional alternative to HTML*; D. MARTIN et al., *Professional XML. Programmer to programmer*, Birmingham 2000; Http://www.wtcm.be/~pool5/news/nl/xml, etc.

A major advantage of XML over ASCII is the use of semantic tags as delimiters. ASCII or Unicode is laid down by ISO (ISO 646, ISO 8859-1, ISO 10646) and is a lasting solution but it does pose a number of problems on the level of user convenience and guarantees of the legal requirements. In an ASCII file in which the records are arranged sequentially, the meaning of certain characters has to be deduced from their position and delimiters are not always clear. Moreover, extensive filter, sort and search activities are not straightforward in large ASCII files. Likewise, conversion to another format may be necessary in that, for reasons of protection of privacy, the electoral register can only be consulted after certain data fields have been deleted to make identification with natural persons impossible. Conversions to more user-friendly formats do not bring us any closer to a platform-independent solution and pose problems when it comes to guaranteeing the reliability of the information. Finally, another problem - along with that of the original - is how to prove with an ASCII file that the legal requirements have been observed.

The XML file consists of a collection of elements that have a certain relation to each other. The structure of a XML file is set out in the DTD (*Document Type Definition*). The DTD describes the relations of the elements and their frequency. Attributes can be assigned to the elements. In the DTD of the electoral register is defined that the XMLfile is made up of the records of voters and that each record comprises three elements: electoral data, personal data and address. These three elements consist in their turn of other elements: electoral data (district, kind, electoral number), personal data (name, date of birth, gender, occupation, nationality), address (street name, house number1, house number2, letterbox number, postal code). The grouping of the thirteen fields that appear in each record is important for efficient data management and fast search activities.

Figure 1: The structure of the XML file is set out in the DTD of the digital electoral register. The electoral register consists of the voter elements (comparable with records), which in their turn are made up of 3 elements. The electoral data, personal data and address elements consist of 3, 5 and 5 elements respectively. The enumeration of elements between brackets and commas indicate that the fields must appear in that order. The "+" sign next to the compulsory voting element signifies that there are several records. A question mark is added to house number 2 and letterbox number 1, indicating that these fields are optional. #PCDATA (*parsed character data*) indicates that the field contains characters that have been parsed.

```
<?xml version="1.0"?>
<!ELEMENT electoral_register (voter+)>
<!ELEMENT voter (voting_details, personal_details, address) >
<!ELEMENT voting_details (district, sort, voting_number)>
        <!ELEMENT district (#PCDATA) >
        <!ELEMENT soort (#PCDATA) >
        <!ELEMENT voting_number (#PCDATA) >
<!ELEMENT personal_details (name, date_of_birth, gender, occupation,
nationality)>
        <!ELEMENT name (#PCDATA) >
        <!ELEMENT date_of_birth (#PCDATA) >
        <!ELEMENT gender (#PCDATA) >
        <!ELEMENT occupation (#PCDATA) >
        <!ELEMENT nationality (#PCDATA) >
<!ELEMENT address (name_of_street, house_number1, house_number2?,
letterbox_number?, postal_code)>
        <!ELEMENT name_of_street (#PCDATA) >
        <!ELEMENT house_number1 (#PCDATA) >
        <!ELEMENT house_number12 (#PCDATA) >
        <!ELEMENT letterbox _umber (#PCDATA) >
        <!ELEMENT postal_code (#PCDATA) >
```

Figure 2: The electoral register in XML format. A XML file consists of a prologue, a body and an epilogue. Though all the prologue and epilogue elements are optional, ideally the prologue should contain a XML declaration (version) and document type declaration (doctype). The document type declaration refers to the DTD to which the XML file must correspond and to the place where the DTD is found. Notice that in the example below the structure corresponds to the DTD (fig. 1) and that the house number2 and letterbox number fields are missing.

```
<?xml version="1.0"?>
<!DOCTYPE electoral_register SYSTEM "electoralregister.dtd">
<electoral_register>
<voter><voting_details><district>18</district><kind></kind><voting_number>00001</voting_number></
voting_details><personal_details><name>VAN ASSCHE ROBERT MARIE
PHILEMON</name><date_of_birth>19580626</date_of_birth><gender>M</gender><occupation>ROOFER</
occupation><nationality>B</nationality></personal_data><address><name_of_street>ABDIJSTRAAT</name
_of_street><house_number1>1</house_number1><postal_code>2040</postal_code></address></voter>
<voter> …</voter>
<voter> …</voter>
…
</ electoral_register>
```

A number of the elements in the electoral register are laid down by law. With the help of a separate application, *the parser*, it is possible to check if it has been drawn up in accordance with the DTD. If this is the case, then parsing produces a validated file. The result of the parsing is reproduced in a logfile, which can be kept as proof that the XML file has been compiled in accordance with the DTD. By validating the XML file, we prove that the digital file does indeed contain the statutorily required entries. By linking the XML file to a data bank we can also ensure that the digital electoral register has been arranged alphabetically or geographically, as laid down by law.

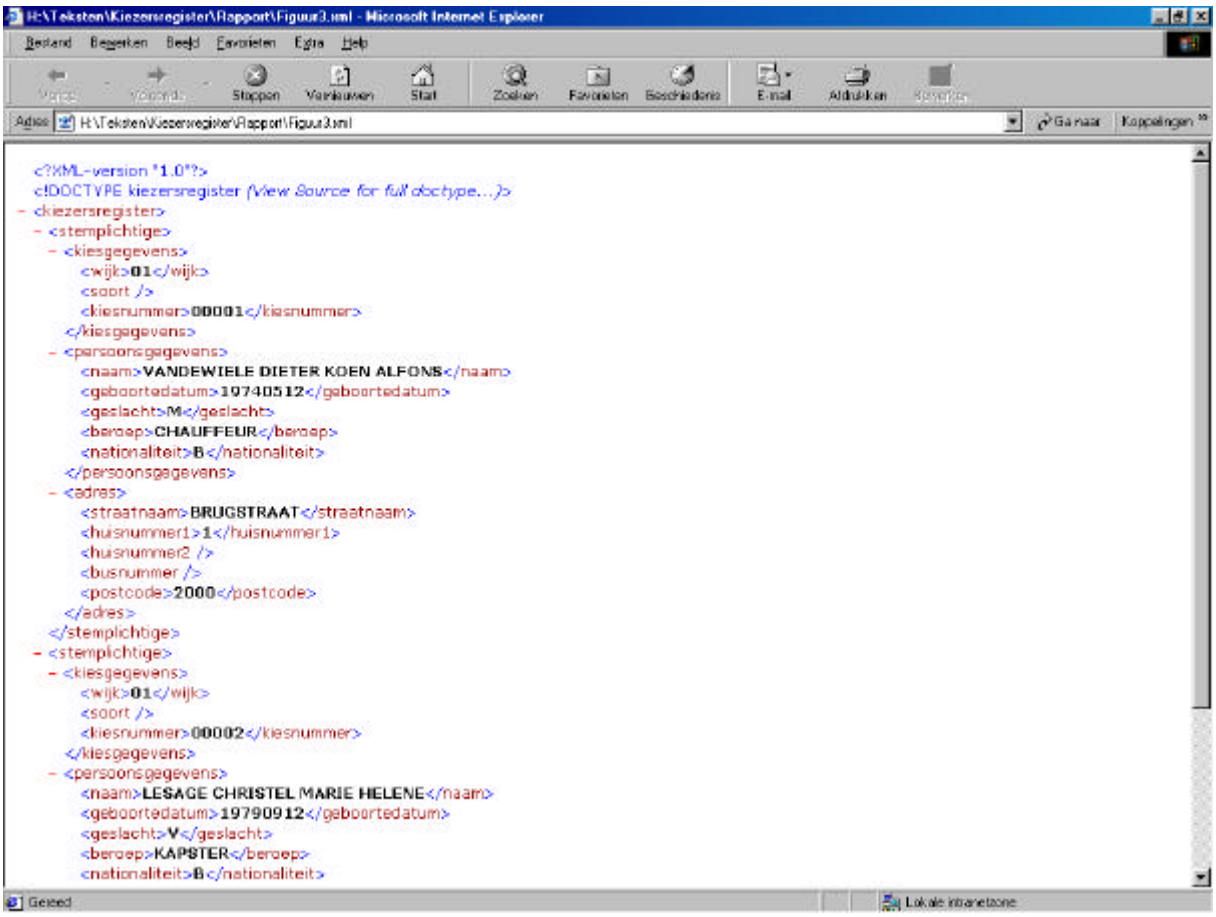
By using a XML Schema instead of a DTD, we can check the content of the fields as well as the structure. This is possible because in the XML Schema not just the structure of the elements, but also the types of data and the valid import possibilities can be defined. For instance, it is possible to check if the date field really does contain a date and, for example, not the name of a person. One might say that a XML Schema combines the advantages of the data structure of a data bank with those of a DTD. Thanks to the import possibilities, the structure of a data bank can easily be converted into a XML Schema.

A XML file can be consulted in various ways. Either you can simply scroll within the file as in an ordinary ASCI file, or the data can be reproduced on the screen by means of an Internet browser (see fig. 3). XML is supported by Internet Explorer 5.0 and Netscape 6.0. Both browsers have an in-built XML parser and retrieve the data by means of a built-in scripting language. The disadvantage of these first two methods is that no modifications can be made (e.g. omitting certain data fields to protect privacy), nor detailed search activities carried out. For large files these are not the most appropriate solutions. Thirdly, style-sheets have been designed for calling up and transforming XML data. By analogy with DSSSL[4] for SGML files, *eXtensible Style-sheet Language (XSL)* has been created for

[4] Document Style Semantics and Specification Language was adopted as standard in 1996 (ISO/IEC 10179) and provides (among other things) a query language for calling up data from the SGML file.

XML. XSL currently has the status of 'candidate recommandation' (Oct. 2000). Several style-sheets can be developed for one and the same file so as to obtain different images. A separate *eXtensible Query Language (XQL)* is currently under development.

Figure 3: The electoral register reproduced as XML file on the screen by Internet Explorer 5.5



XML was approved by the W3C in February 1998 and has the status of 'recommandation', which is the W3C term for standard. The major software producers (SoftQuad, Adobe, IBM, HP, Microsoft, Lockheed Martin, NCSA, Novell and Sun Microsystems) are represented in the XML work group, as are universities (Boston, Oxford, Illinois, Waterloo) and government bodies. So we can expect the software producers to make their applications compatible with XML or to incorporate XML into their products, and indeed this is already happening[5]. Several XML parsers are already available; as a rule, these can be downloaded from the Internet free of charge. XML is as yet not an ISO standard[6]. In fact, by ISO norms, XML is not yet an official standard but more a *de facto* standard. Yet XML is classified higher in the hierarchy of standards than other *de facto* standard file formats such as Word documents or PDF files. It owes that status to its platform and vendor independence. At least as much can of

[5] Microsoft Office 2000 uses XML as the format for exchanging data. In the medical world various applications are used for storing patients' medical files as XML files.

[6] See: E. DUMBILL, *Should XML become a "real" standard?*, Nov. 2000 (Http://www.xml.com/pub/a/2000/11/devcon/standards.html)

course be achieved with the W3C standard as with an ISO norm: the major hardware and software producers take XML into account when planning future developments. Moreover, it is highly unlikely that XML will exist for ever. Information Technology will continue to develop and it is inconceivable that there will never be a better alternative to XML. Even if we take the worst scenario imaginable (the disappearance of XML altogether), then we still have a platform and vendor-independent file in which tags give a meaning to the content of the fields. Then XML has the great advantage that all its syntax characters are ASCII characters. And there will certainly always be some considerable compatibility with SGML or other new standards.

At the moment this applies to a much lesser extent to XML Schemas. XML Schemas only became a 'candidate recommandation' on October 28th 2000 and there are only a limited number of parsers available that are suitable for XML Schemas. For the digital archiving of the electoral register of the City of Antwerp, a DTD rather than a XML Schema was drawn up. With the digital archiving of the electoral register of the City of Antwerp, we could not use the DTD model in its entirety but had to alter it on a number of fronts. This is largely the result of the fact that the mainframe file containing the electoral register was converted retroactively into XML. At mainframe level each record consists of thirteen sequential fields. The task of clustering the thirteen fields in three elements was not done. The sequence of the thirteen fields differs radically from those of the DTD model and it would be a lot of work to change the order again. The beginning and end tags have to be added between the fields, whereby the beginning position of a field must constantly be shifted by the same number of positions as the total number of characters of all the beginning and end tags between the zero position of the record and the first position of the field. To ensure this operation runs smoothly, we opted for tags with a fixed field length. Because this limited the possibilities for giving tags a semantic meaning, we chose a code as the tag. The meaning of the code is set out in the list of metadata. For the same reason the house number2 and letterbox number elements are not optional and they appear in every record[7].

With the electoral register as an XML file, we have a digital file that can be permanently and easily consulted and that satisfies the legal requirements. Finally, we must ensure that it is archived as a reliable file. The file must be authentic and intact. There must be absolutely no doubt about the provenance and the accuracy of the information. The provenance of the electoral register is obvious. The electoral register is an administrative document of the Antwerp municipal authorities and is drawn up by its administration in conjunction with its computer service center. The archived example is preserved and managed by the Antwerp City archives. As long as the electoral register can only be consulted in the reading room of the archival service, we need take no special measures to prove its provenance. However, when it comes to digital files that can be consulted, for example, via the Internet, things are rather different, but this problem does not present itself in the case of the electoral register.

The Administration is responsible for the accuracy of the information contained in the data file. The archival services must ensure that the contents of the file are not altered once it has been transferred. The measures we can take for this, depend on the way the file is made available to researchers. If the file is consulted on CD-ROM, then we prefer to use CD-ROMS that cannot be rewritten or we can equip the reading room with machines without a CD-writer. If the file is on the file server of the archival service and it can be consulted on the intranet, then we can place the files behind a fire wall or lay down the user rights of visitors to the reading room in such a way that they are not

[7] The DTD that was used by the Antwerp City Archives can be consulted on http://www.antwerpen.be/david. (Follow the button XML-DTD).

empowered to introduce changes (e.g. read only) or can only use certain (parts of) applications (e.g. 'save' or 'save as').

It is possible to find out if a digital file has been altered by hashing. Hashing is a technique used to translate the contents of a file (a combination of bits) into a unique code by means of a mathematical algorithm. If just one bit of digital data is altered, then a totally different hashing code is obtained assuming the same algorithm is used. The archive file can be hashed during transfer to the archival service. The algorithm used and the hashing code are kept separate from the archived file, for example in a data bank which is specially reserved for this purpose. When the archivist wants to check the reliability of the information, then he again calculates the hashing code on the basis of the algorithm used. If the result of the second hashing operation is identical to the first hashing code, then he can be certain that the file has not been tampered with.

B. THE STORAGE MEDIUM

A CD-ROM was chosen as the storage medium for the digital electoral register. A CD-ROM meets the specified requirements in terms of a durable medium with sufficient storage capacity. The standard norms ISO 9660 and ISO 10149 cover CD-ROMs. The ISO 9660 standard relates to the volume, the folder and file structure and the name given to the file, while ISO 10149 governs the way CD-ROMs are recorded. The application of the two norms means that the CD-ROM containing the electoral register can be consulted on all the platforms that support these standards[8].

When storing CD-ROMs we must ensure their durability and that they continue to be usable; factors related to their life span and commercial support. The classic CD-ROMs are made of polycarbonate and metal and so have a limited life span (10 to 20 years), even when stored in the ideal climatic conditions (approx. 20 °C and 40 % relative humidity). Work is currently being carried out on CD-ROMs made of materials with a longer life span or materials that are more resistant to fluctuations in temperature and humidity (glass, nickel, iridium, etc.). The difficulty with CD-ROMs, however, is that deterioration or damage usually occurs without warning. This problem can be offset by regularly checking the number of faults on the CD-ROM and by keeping back-up copies. Digital files do have the advantage that they can easily be transferred to another medium.

IX. EXTRAPOLATION

The electoral register is an example of a static document consisting of fixed data fields; only its content and structure are important. So the way we archive the electoral register digitally, can also be

[8] http://www.uni-regensburg.de/EDV/IO_Server/scripts/manual/node17,
http://hydra.mpistuttgart.mpg.de/zwe/dv/iso9660.html,
http://www.ccs.neu.edu/home/bchafy/cdb/info/iso9660.txt,
http://www.openvms.compaq.com:8000/72final/4506/4506pro 001.html,
http://www.oreilly.com/reference/dictionary/terms/C/Compact_Disc_Read-Only_Memory.htm

applied to files with similar properties. Typical examples of such digital files are data banks and text files consisting of fixed fields. Illustrations cannot as yet be included in XML files, though a reference and the relevant viewer which has to be opened, can.

Antwerp City Archive applied this archiving strategy not only to the electoral register, but also used it for archiving data from the municipal register. By converting the application used to keep the population data up to date (Dec. 1999), the records of the person which had died or left the city before 1983 were not transferred. These records had to be archived. It contains the personal details of approximately 80,000 people. 175 fields, broken down into a number of segments, were kept for each record[9].

X. CONCLUSION

The legal analysis shows that the electoral register does qualify for digital archiving. The current legislation and regulations contain no obstacles that pose a threat to digital archiving. The legal requirements with regard to form can be carried over into a digital environment and the legislator provides for the digital production of the electoral register. So we can confirm that the electoral register can be archived digitally. The requirements with regard to form, which the electoral register has to meet, are limited: a number of obligatory entries, continuous numbering, drawn up per district of the municipality, alphabetically or geographically according to the streets. There was no problem meeting these obligations in the digital electoral register.

The relevance of the permanent preservation of the electoral register can be questioned. The electoral register is a derivative of the municipal register and has no archival excess value whatsoever in digital form. The statutory requirement which rules that the digital electoral register be drawn up per district of the municipality, alphabetically or geographically, makes little sense in a digital context.

XML was the chosen file format for archiving. The descriptive mark-up language XML is platform and vendor-independent and meets the requirements governing usability, reliability, durability and access. With XML we can archive the electoral register without loss of data and in such a way that the legal formal requirements are met. By parsing the XML file using a DTD or XML Schema, we validate the structure and content of the digital electoral register and show that the legal requirements have been met. The authenticity of the digital electoral register presents no problems because the electoral register is only consulted internally, namely in the buildings of the archive institute. When making it accessible to third parties, we take measures to prevent the file being altered. With the help of the hashing technique, we can prove that the file has not been altered whilst the responsibility of the archival service. In this way we guarantee the integrity of the electoral register.

Provided a number of modifications are made, the model strategy can be put into practice. The fact that we have to deviate from our DTD model on a number of points is a consequence of the retroactive conversion to XML. This can be avoided in future by archiving the electoral register directly in XML.

---

[9] The DTD used for this, can be consulted on http://www.antwerpen.be/david.