

An administration without electronic office documents isn't conceivable in this time and age. Undoubtedly electronic offices offer many advantages but, the management and record-keeping of electronic documents often poses a number of practical problems: it is difficult to re-find an electronic document, computerfiles which are just a couple of years old can no longer be read, the same documents are present in an electronic and a paper form, files are deleted to release hard disks or crammed mailboxes etc. These problems are probably recognizable, but how can you avoid them?

1. Electronic documents in the archive?

Many documents within the administration have archival value and need to be preserved for a long time. These documents have an administrative or judicial value or can be used as a corporate memory. For a number of document categories, archiving is a legal obligation and a minimal retention period is set for each of them. Besides, for Belgian public institutions it is decreed that all documents are public documents and are therefore subject to the public records act.

Public institutions are obliged by law to keep their archives in a well-ordered and accessible state. Furthermore, the approval of the General Archivist of the Belgian National Archives or his authorized representative is required when destroying records of public institution or a Social Service Department, according to the archival legislation. These legal obligations apply for paper as well as for electronic documents. Both kinds of documents can have the status of an administrative or archival document. It doesn't matter whether it is a regular letter, an e-mail or a database. The obligation to archive implies, among others, that electronic documents have to be traceable and readable.

From an archival point of view, it is best to save records in their original form. What is born digitaly (on paper), should be archived digitaly (on paper). E-mails and websites are examples of records that are archived digitaly. The principle is that when the same record exists on paper as well as digitaly, the original version is kept. In the case of a scanned letter, this means that the paper letter is kept in the archive. It is possible to differ from this, but only when the digitalisation complies to certain quality demands. And even that case, the approval of the General Archivist of the Belgian National Archives is necessary.

2. Organizing an electronic office

2.1. An electronic classification system

Take your time to devise an electronic classification system to ensure that the documents can be re-found easily after their creation and use. A well thought-out and logically designed classification system is very important, as it is for paper documents. Although full text searches are possible on electronic documents, these searches often lack efficiency.

In an electronic context, the classification scheme is designed as a folder structure. Such a folder structure can be compared to a tree structure which branches off. It is recommended to design a logical and surveyable folder structure (for every agency) in which the documents can be classified by case or subject. This way an electronic file is created in which textfiles, spreadsheets etc. can be classified together. Also, e-mails should be kept on the common serverspace. E-mails shouldn't be stored in the e-mailsystem but exported to the matching folder after receipt or mailing. This way you avoid scattering electronic information all over the mailsystem and the folder structure. The "record-keeping" possibilities of e-mailsystems shouldn't be used.

Some rules to keep in mind:

- make clear-cut agreements within the organisation on which documents are kept on the common serverspace,
the personal folder and possibly on the local hard disk.
- avoid storing important documents on the local hard disk. Files on the common serverspace are better secured, back-ups are made and they are available for colleagues, etc....
- base the classification of the folder structure on the workprocesses and tasks and activities of the department.
- give the folders a clear-cut, unique and meaningful name.
- include a classification code in the foldername. By placing the classification code in front of the foldername, you avoid the disadvantages of an alphabetical sorting and the structure is build up from general to specific. The classification code can be used to refer to a case or subject.
- if possible, base the classification on the classification system used for paper documents. Use the same classification code for paper and electronic files so the relation is obvious.

These rules are easy to apply within the common operatingsystems. These are the first steps towards an organized document- and information management. The electronic records stay usable for the organisation: searching for information takes up less time, it's easier for colleagues to refind documents, making duplicates is avoided, the latest versions can easily be found, etc. The functionalities of a common operatingsystem can be too limited for certain categories of documents. More sophisticated document management systems can be appropriate for important or documents with sensitive information because of the tighter security, version control or automatic logging. A comparable document organisation can be applied within the document management systems.

2.2 Meaningful filenames

Within this kind of electronic classification scheme, the admission of meaningful filenames is important. Filenames identify the electronic office documents and have to lead the user to the desired information. It's advisable to follow a number of rules when assigning filenames, such as:

- do not repeat the foldername in the filename.
- assign clear-cut, unique and meaningful names to the files. Indicate the kind of document which is involved, the subject, who the sender/receiver is, the date, the versionnumber etc.
- make sure that the filename and the documentname or- title match
- keep the extension of the application in the filename.

3. Durable fileformats

The force of electronic information can be found, among others, in the re-usability. Therefore, electronic information is stored in a certain kind of fileformats. The choice of the fileformat in which the document with archival value is stored, is important. After all, hard- and software age fast and become obsolete after a while. The retentionperiod of electronic documents often outlasts the life cycle of specifique hard- and software. Because of these reasons, it is recommended to use as many standards as possible for electronic records which have a long retentionperiod and to avoid storing them in application or proprietary formats.

Specific application formats are usually tied to one particular manufacturer, computer application or even version. The files need to be converted to computerprogrammes that are common at the moment when

changes occur in the available software or support. Such conversion operations are expensive, take up a lot of time and need to be repeated regularly. Consequently, it isn't advisable to archive documents in a fileformat such as, for instance, MS-Office. An exception is PDF from ADOBE. The specification of the PDF-format is freely available so each developer can write a PDF-compatible programme. This isn't the case for the MS-Office fileformats. Presently the official standardisation of a PDF-subset for record-keeping purposes (PDF/A) is being worked-out. The definitive standardisation is expected by the end of 2003.

It is advisable to aim for as much independency from specific computer programmes as possible when archiving, and to choose fileformats which can be read by several kind of applications. Such examples are Unicode/ASCII, XML, TIFF, MPEG, etc. Documents which are stored in a standardised fileformat do not need to be converted when a computerapplication is no longer available. The composition of these formats is documented in the standard and their management isn't controlled by one commercial company, but by a standardisation authority.

Appropriate record-keepingformats

text:	Unicode (ASCII), XML and PDF(/A)
image:	raster: standard TIFF for mastercopies (no-compression, high resolution), JPEG voor safetycopies or distribution vector: CGM, EPS, DXF, SVG
sound:	compressionless WAV (PCM-coding)
video:	MPEG

It is recommended to keep office documents in an appropriate record-keepingformat as soon as possible, if possible from the moment of creation. OpenOffice, the opensource alternative for MS-Office, keeps the documents directly as XML-files which can be opened with any other application (for instance a webbrower or texteditor). Though this isn't always feasible. When choosing a storageformat one needs to consider the possible future re-use. The Antwerpian city administration keeps its e-mails temporary as msg-files, so that the exported e-mail can still be answered or forwarded. The mail is migrated to XML before being archived.

4. Write to tape or cd

To keep hard disks and networks performant, it is necessary to remove closed files and to store them on tape or optical disc. Although not every tape or optical disc is suitable for archiving purposes. Standards apply here too. Making back-up tapes for archiving is not an option because back-ups are hard- and software dependent.

4.1 Archiving on tape

- Use cartridges (1 reel) or cassettes (2 reels). Avoid open reels.
- Choose an interchangeable, non-manufacturer dependent physical cassette- or cartridgeformat. Only use standardised cassettes or cartridges. Check in advance the number of times the carrier can be used (how many insert & eject-cycles or passes?)
- Use a tape with MP (metal particulate) or barium ferrietpartickels on a basis of polyester
- Choose a technology based on the linear recordingmethod (*longitudinal* or *serpentine*): for instance DLT

or LTO.

- Make labeled tapes. Apply the standards ANSI INCITS 27-1987 (R1998) or ISO-1001
- Fill out the volumelabel of the tape
- Document the tape. Keep information on: fixed or variable-length blocks, block size and records length, density, filesystem, used software applications (+ version), encoding

4.2 Archiving on CD

The cd-rom is probably the best known optical carrier for electronic information. Cd-r's are cheaper and most computers are equipped with a cd-rom station. It will be possible to use dvd's in the near future, but for the moment the standardisation of the recordable DVD isn't finished yet. In the mean time, archiving on cd gives the best guarantee.

Make quality cd's:

- first make an image of the cd-rom on the local hard disk of the computer which is equipped with a cd-burner. Do not burn cd's over a networkconnection.
- provide enough free memory and diskpace. Close other computerprogrammes (such as virus control programmes and screenscreens)
- use the 'disk-at-once' method
- avoid multi session cd's and close it immediately after writing.
- use different brands of blanc cd's for the master- and safetycopy
- use 650 MB cd's and write up to a maximum of 630 MB (no overburn)
- apply the official standard ISO-9660
- foldernames shouldn't contain more than 31 characters, filenames 8 (level 1) or 30 (level 2)
- only use the characters A-Z, 0-9, _
- don't create more than 8 levels
- do not use Joliet (Windows) or Rock Ridge (Unix) extensions
- apply the official standard ISO-10149; choose: CD mode 1

5. More info

More information on the DAVID-project can be found on the website; <http://www.antwerpen.be/david>. DAVID ('*Digital Archiving in Flemish Institutions and Services*'). is the first Flemish researchproject on electronic record-keeping. The project tries to answer to the demand for practical information on electronic record-keeping with the serie *Digital Archiving:guldeline & aDvice*. Practical information on the preservation of the electoral register, the population register, e-mails, websites and GIS is also available on the website. The project plans a manual on electronic record-keeping by the end of 2003.