

### E-mail omzetten naar XML

Bestandsformaten zijn niet bestand tegen de tand des tijds. Software verouderd, raakt buiten gebruik en wordt vervangen door toepassingen die nieuwe formaten ondersteunen. Regelmatig migreren naar een nieuw bestandsformaat is onvermijdelijk.

Een geschikt **archiveringsformaat** is open en niet-applicatiegebonden. Doordat het goed gedocumenteerd is, kan ook na lange tijd nog achterhaald worden hoe de bestanden te lezen. Voor e-mail is **XML** het meest geschikte archiveringsformaat. XML laat toe extra informatie aan een archiefdocument toe te voegen door het omsluiten van de verschillende gegevensvelden met **XML tags**.

```
<email>
...
<correspondent>
<name>eDAVID</name>
<emailAddress>info@edavid.be</emailAddress>
</correspondent>
...
</email>
```

Een **XML schema legt** de structuur van een XML document vast. Op de eDAVID-site staan **2 XML schema's** voor het archiveren van e-mail ter beschikking: *email.xsd* en *addressbook.xsd*. Deze schema's kunnen zo gebruikt of worden aangepast aan de eigen noden.

Het omzetten naar XML is niet nog niet automatisch mogelijk met de courante e-mailsystemen. Wel bestaan er hiervoor speciale tools en is het voor de meeste e-mailcliënts ook mogelijk om een toepassing op maat te ontwikkelen. Ook eDAVID ontwikkelt nieuwe tools om verschillende bestandsformaten naar XML om te zetten.

### Archiveringsformaat voor bijlagen

Omdat er een grote diversiteit is in de aard van bijlagen zijn ook de geschikte archiveringsformaten divers. In "Digitaal Archiveren: rIchtlijn & aDvies, nr. 4, Standaarden voor bestandsformaten" lees je hierover meer.

### Kwaliteitszorg voor digitale archiefdocumenten

Na het vastleggen van de essentiële componenten en contextgegevens en het omzetten naar een geschikt archiveringsformaat zijn de archiefdocumenten **bewaarklaar**. Hiermee is de laatste stap echter niet gezet. Digitaal archiveren op lange termijn is een proces van **continue kwaliteitszorg**. Houd bij het uittekenen van je bewaarstrategie ondermeer rekening met volgende elementen:

1. start met de keuze van een **geschikte drager** voor bewaring op lange termijn. Meer informatie o.a. in: "Digitaal Archiveren: rIchtlijn & aDvies, nr.2: Optische dragers." en "Digitaal Archiveren: rIchtlijn & aDvies, nr.6: Duurzame magnetische dragers"
2. gebruik een standaard **bestandssysteem**.
3. zorg voor **veiligheidskopieën** voor het geval er toch iets misloopt.
4. voorzie regelmatig **kwaliteitscontroles**:
  - dragers van digitale informatie (harde schijven, dvd's, cd's, enz.) zijn kwetsbaar. Het materiaal zal degenereren en gegevens kunnen niet meer foutloos worden ingelezen.
  - ook op het niveau van de bits en bytes kunnen fouten ontstaan.Zorg daarom voor systematische controles op beide niveau's om na te gaan of de informatie nog intact is, indien niet kan de informatie dankzij de veiligheidskopieën teruggeplaatst worden.
5. volg de **technologische ontwikkelingen** continu op, na verloop van tijd:
  - zullen migraties naar nieuwe geschikte bestandsformaten nodig zijn.
  - komen er misschien nieuwe, meer geschikte dragers op de markt.

### eDAVID onderzoek

Het digitaal depot project onderzoekt de kwaliteits-, functionele en juridische vereisten waaraan een depot moet voldoen en legt hiermee een basis voor het uitbouwen van een bewaarservice van digitale depots. Alle informatie over het Digitaal Depot project vind je op de project-site: <http://www.edavid.be/digitaaldepot>

De in deze folder vermelde - en nog vele andere - eDAVID publicaties zijn online terug te vinden: <http://www.edavid.be> Nog vragen? [info@edavid.be](mailto:info@edavid.be)



**Waardevolle e-mails bedreigd**

E-mail is inmiddels onmisbaar bij communicatie zowel thuis als op kantoor. E-mail behoort soms ook tot het cultureel erfgoed: waar bijvoorbeeld vroeger brieven van auteurs hun weg naar een archief of bibliotheek vonden, verdwijnen de e-mails die hun plaats innemen vaak ongemerkt.

Net als andere digitale documenten is e-mail om diverse redenen kwetsbaar en vluchtig (zie eDAVID, *Folder 1: Bewaren van Digitaal Cultureel Erfgoed in een Digitaal Depot*, september 2006). Daarom moeten tijdig stappen genomen worden om ze te bewaren voor de toekomst.

**4 stappen voor bewaring**

De tips die het Expertisecentrum DAVID (eDAVID) in deze folder aanreikt, helpen éénieder bij het organiseren, bewaarklaar maken en duurzaam bewaren van e-mail en bijlagen.

4 eenvoudige stappen volstaan:

- 1. vastleggen van de essentiële componenten**
- 2. situeren in de context**
- 3. migratie naar een archiveringsformaat**
- 4. duurzame bewaring**

**Essentieel: inhoud, structuur en transmissiegegevens**

Wanneer we digitale documenten duurzaam en efficiënt willen bewaren, is het belangrijk om allereerst de componenten te identificeren die onontbeerlijk zijn.

**Inhoud en structuur** worden algemeen als essentiële componenten van een e-mailbericht beschouwd. Onder inhoud vallen onderwerp en het bericht zelf, maar ook eventuele bijlagen.

**Opmaak** is niet altijd essentieel voor een e-mail. De beslissing om al dan niet ook *look & feel* te bewaren is afhankelijk van de appreciatie van de archiefbeheerder.

Ook de **transmissiegegevens** vertellen veel over een e-mail. Deze gegevens vind je terug in de e-mailheader zoals: de naam en het e-mailadres van de afzender en van elke geadresseerde (Aan,CC, BCC), de datum en het tijdstip van verzending en van ontvangst en het aantal bijlagen. Transmissiegegevens zijn essentieel om een authentieke e-mail te archiveren.

```
Content-Type: message/rfc822; name="eDAVID-
Nieuwsbrief"
Content-Transfer-Encoding: 8bit
Delivered-To: info@edavid.be
Date: Wed, 28 Jun 2006 11:02:55 +0200
User-Agent: Thunderbird 1.5.0.4
Content-Transfer-Encoding: 8bit
Content-Type: text/html; charset="utf-8"
```

**Dynamische gegevens tijdig vastleggen**

Sommige belangrijke gegevens zijn **niet (volledig) zichtbaar**. Afzonderlijke e-mailadressen of volledige verzendlijsten worden weergegeven als *nicknames* die je zelf aanmaakte in het adresboek.

Daarbij komt dat adresboeken en verzendlijsten **geen statische objecten** zijn. Gegevens worden gewijzigd, toegevoegd of geschrapt. Na verloop van tijd is het niet meer te achterhalen welke samenstelling een verzendlijst had op het moment dat je een e-mail verzond. De samenstelling van de verzendlijst wordt daarom best samen met de e-mail bewaard. Ook bij andere gegevens zoals datum en tijd loop je het risico dat ze door de e-mailcliënt worden bijgewerkt.

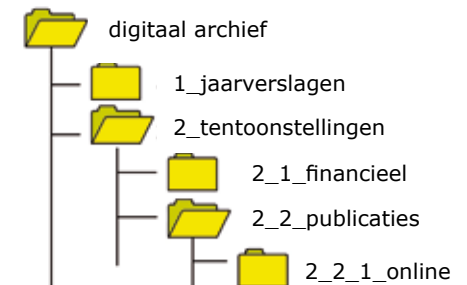
Daarom worden er best maatregelen genomen om dynamische gegevens tijdig vast te leggen. Dit kan door middel van programmatorische ingrepen en/of export.

**Mappenstructuur zorgt voor context**

Een afzonderlijke e-mail is na verloop van tijd soms nog moeilijk te begrijpen door het ontbreken van verwijzingen naar andere e-mails, bijlagen en documenten. Het is daarom belangrijk zoveel mogelijk van deze informatie te bewaren zodat de documenten ook later nog in hun **context** kunnen worden gelezen. Een eenvoudige maar efficiënte oplossing is het beheren van je e-mails via een **mappenstructuur**. Door de groepering in mappen en de relaties die zo tussen documenten gelegd worden, verschaft je belangrijke contextinformatie.

Concrete richtlijnen, tips en voorbeelden voor het uitwerken van een digitaal klassement lees je in: "*Digitaal Archiveren: rIchtlijn & aDvies, nr. 3, Mappenstructuur en bestandsnamen voor digitale documenten, januari 2002.*"

Binnen de e-mailclient is het mogelijk mappen offline te zetten (als lokale kopie i.p.v. op de mailserver). Het bestandsformaat is afhankelijk van de e-mailclient: .mbox voor *Eudora* en *Thunderbird*, .pst voor *Outlook*, .dbx voor *Outlook Express*, enz.

**Integratie in een digitaal klassement**

Het bewaren van e-mail via de mappenstructuur van een specifieke e-mailcliënt is slechts een tijdelijke oplossing. Voor bewaring op langere termijn integreer je e-mail, bijlagen en andere gerelateerde digitale documenten die niet via e-mail werden verzonden in één **digitaal klassement**. Met een dergelijk klassement is het terugvinden en herbruiken van bestaande informatie eenvoudig.

In veel e-mailcliënts is het exporteren van mappen een ingebouwde functie of kan er een plug-in voor geïnstalleerd worden. Afzonderlijke e-mails kunnen in dit klassement worden bewaard in verschillende formaten, afhankelijk van het clientprogramma: .eml, .msg, .txt, enz.