

# Wat zijn conserveringsmetadata?

Henk Vanstappen, PACKED vzw

Om digitale bestanden duurzaam te bewaren en toegankelijk te houden, wordt beroep gedaan op een 'digitaal depot'. Eén van de belangrijkste vereisten voor een duurzame digitale bewaring is de documentatie van conserveringsmetadata. Wat zijn preservation metadata, waarvoor heb je ze nodig en om welke gegevens gaat het precies?

## PRESERVATION METADATA?

Om duurzame toegankelijkheid van digitale bestanden te verzekeren, moet je verschillende vragen kunnen beantwoorden:

- Kan het bestand nog worden gelezen?
- Wordt het bestand op de juiste manier weergegeven?
- Hoe moet ik het digitale object precies interpreteren? Wat is de context van het bestand?
- Is het digitale object wat het beweert te zijn? Werd het niet niet bewust of onbewust gewijzigd?
- Hoe heet het digitale object en waar kan ik het vinden?

Conserveringsmetadata zijn de gegevens die je nodig hebt om deze vragen te beantwoorden. Enkele voor de hand liggende voorbeelden zijn:

- een unieke identificatiecode waarmee een digitaal object wordt aangeduid, zoals een persistente URI;
- de bewaarplaats van het bestand, zoals het netwerkpad of een URL;
- informatie om te bepalen of een object werd gewijzigd, zoals een MD5 checksum.
- een precieze aanduiding van het formaat van het digitaal bestand.

Daarnaast worden vaak ook eigenschappen vastgelegd die nodig zijn om het conserveringsbeleid te bepalen: welke eigenschappen van het bestand moeten precies bewaard blijven? Dergelijke metadata noemt men de 'essentiële eigenschappen' (significant properties) van een bestand. Wanneer je bijvoorbeeld van een Word-bestand enkel de inhoud zelf wil bewaren, zonder de opmaak, vormt de tekst de (enige) essentiële eigenschap.

Conserveringsmetadata kunnen rechtstreeks betrekking hebben op het digitale object, maar ook op de acties die het heeft ondergaan en op



Conserveringsmetadata bevatten alle gegevens om bestanden bruikbaar te houden. Foto: NASA.

de personen of toepassingen die deze acties hebben uitgevoerd, zoals de datum waarop het bestand in het eDepot werd opgenomen.

## PREMIS

Om conserveringsmetadata uitwisselbaar te maken, werd een XML-schema ontwikkeld: PREMIS: datadictionary for preservation metadata.<sup>1</sup> PREMIS onderscheidt meer dan honderd (!) verschillende gegevens die van belang kunnen zijn om een duurzame bewaring en leesbaarheid van een bestand te garanderen.

Vanzelfsprekend worden niet al deze elementen overal gebruikt en toegepast. PREMIS is meer een staalkaart van metadata-elementen, waaruit je een overwogen keuze kan maken.

## DUURZAME METADATA

Belangrijk is dat conserveringsmetadata zelf ook duurzaam bewaard worden. Dat betekent vooreerst dat ze in een neutraal formaat moeten worden opgeslagen, zoals XML. Verder moet ook verzekerd worden dat de relatie met het digitaal object zelf bewaard blijft. Dat kan door de verwijzing aan de hand van een (persistente) identifier, of door het bestand en de metadata als een geheel te bewaren, bijvoorbeeld in de vorm van een Bagit.<sup>2</sup> ■■

- 1 PREMIS Editorial Committee. (2008). PREMIS Data Dictionary for Preservation Metadata, Version 2.0. url: <http://www.oclc.org/research/projects/pmwg/premis-final.pdf>
- 2 Network Working Group. (2014). The BagIt File Packaging Format (V0.97) url: <http://www.digitalpreservation.gov/documents/bagit-spec.pdf>

> Dit artikel werd bezorgd door PACKED vzw dat als expertisecentrum digitaal erfgoed kennis, ervaring en deskundigheid omtrent digitaal erfgoed centraliseert en verspreidt. Vragen voor PACKED vzw zijn welkom via [info@packed.be](mailto:info@packed.be).