Duurzame toegang tot onderzoeksdata nog zeer fragmentarisch

INGE ANGEVAARE

Duurzaam archiveren en toegankelijk houden van wetenschappelijke data is nog steeds meer uitzondering dan regel, zo blijkt uit het onlangs verschenen interimrapport *Toekomst voor ons digitaal geheugen: duurzame toegang tot informatie in Nederland*. Dit in tegenstelling tot wetenschappelijke publicaties, die via de digitale archieven van de universiteiten duurzaam worden opgeslagen in het e-Depot van de Koninklijke Bibliotheek.

In de Nationale Verkenning Digitale Duurzaamheid, die ten grondslag ligt aan het rapport, werd het onderscheid tussen digitale publicaties en onderzoeksdata al snel duidelijk. Voor de eerste categorie bestond er in het analoge tijdperk een gevestigde praktijk. De informatie werd door de auteur naar een uitgever gestuurd, na *peer review* gepubliceerd in een vaktijdschrift, en vervolgens duurzaam opgeslagen in (nationale) bibliotheken. Die praktijk kon in het digitale tijdperk grotendeels worden voortgezet.

Complicerende factoren

Digitale onderzoeks*data* vormen een heel nieuwe categorie, waarvan de verkenning uitwees dat voor duurzame opslag en hergebruik nog maar incidenteel voorzieningen zijn getroffen. Daar zijn een aantal oorzaken voor. Onderzoeksdata zijn er in ontelbare vormen en formaten, en ze bevinden zich overal en nergens – op de eigen PC van een onderzoeker, op servers van instellingen, in grote internationale datacentra. Ze bewegen zich tussen al die netwerken en worden steeds opnieuw bewerkt. Dat alles maakt het lastig om grip te krijgen op die informatiestromen. Zeker internationaal is het immers niet gemakkelijk om partijen aan te wijzen die verantwoordelijk zijn voor de organisatie van duurzame toegankelijkheid én voor de financiering daarvan.

Een andere complicerende factor is het feit dat de onderzoeker die de data produceert het belang van duurzame toegankelijkheid vaak niet als het zijne ervaart. Uit onderzoek in het kader van andere projecten, zoals PARSE.insight, blijkt ook dat onderzoekers wel data van anderen willen gebruiken, maar zelf veel minder bereid zijn om hun data beschikbaar te stellen aan derden. Ze vrezen dat anderen misbruik zullen maken van het door hen verrichte werk en ervaren de extra inspanning die van hen wordt gevraagd als lastig en tijdrovend. Het belang van de samenleving, die als geheel wel baat heeft bij het hergebruiken van onderzoeksdata, staat verder af van de individuele onderzoeker.

Selectie nodig

En dan is er de grote vraag wat wel en niet moet worden opgeslagen voor hergebruik. Bij wetenschappelijke publicaties fungeert het uitgeversbedrijf met behulp van de *peer review* als selectiemechanisme. Een dergelijk mechanisme ontbreekt voor onderzoeksdata, terwijl de hoeveelheden veel groter zijn en de verschijningsvormen (ook in technische zin) veel gevarieerder en vaak ook ingewikkelder dan bij publicaties, waar pdf de boventoon voert. Er zijn ook grote verschillen tussen disciplines: daar waar onderzoek onmogelijk is zonder longitudinale data, bijvoorbeeld in de meteorologie, is duurzame opslag al snel tot stand gekomen. Ook wettelijke regels kunnen helpen, zoals in de archeologie waar het Verdrag van Malta dat deponeren van onderzoeksgegevens verplicht stelde. Archeologische data vormen de grootste categorie in de archieven van Data Archiving and Networked Services (DANS).

Buiten DANS zijn in Nederland maar enkele wetenschappelijke data-archieven bekend die echt beleid en organisatie hebben afgestemd op toegankelijkheid voor de lange termijn. Voorbeelden zijn het Instituut voor Nederlandse Lexicologie, het Max-Planck-Institute for Psycholinguistics, het Internationaal Instituut voor Sociale Geschiedenis en het KNMI. Een datacentrum van de drie technische universiteiten (3TU) verkeert in de opbouwfase. Verreweg de meeste onderzoeksdata vallen momenteel nog buiten enig duurzaamheidsregime.

Het ontwikkelen van selectiemechanismes is een belangrijke richting waarin volgens het rapport gewerkt zou moeten worden om duurzame toegankelijkheid te bevorderen. Ook zouden onderzoekers die hun data ter beschikking stellen op de een of andere manier beloond moeten worden voor hun inspanningen. Financiers zouden deponeren verplicht kunnen stellen, zoals NWO met zijn programma Middelgroot al doet. Datacentra, ten slotte, kunnen en moeten onderzoekers helpen bij het organiseren en beheren van hun data, opdat die na afloop van het onderzoek geschikt zijn voor hergebruik.

www.ncdd.nl

Er zit toekomst in het verleden

MARJOLIJN KRUIP

Middeleeuwse pelgrimsroutes op de kaart gezet

Op de afdeling Kunstgeschiedenis van de Radboud Universiteit wordt al jaren onderzoek verricht naar pelgrimage in de late middeleeuwen. Dit onderzoek heeft een nieuwe impuls gekregen door de ontwikkeling van een internettoepassing met geo-technologie.

Pelgrims die een bedevaartsoord hadden bereikt, kochten daar in de late middeleeuwen vaak insignes of ampullen. Deze religieuze objecten herinnerden de gelovigen door de afbeelding aan het bezochte heiligdom; daarnaast boden ze bescherming, weerden ze het kwaad af en brachten ze geluk. In de onderzoeksgroep Kunera, onder leiding van professor Jos Koldeweij, wordt al ruim tien jaar gewerkt aan een database van zulke laatmiddeleeuwse pelgrimssouvenirs.

Souvenirs die in de ene plaats werden gemaakt en in de andere als bodemvondst opduiken, geven inzicht in het reisgedrag van laatmiddeleeuwse pelgrims. Daarom, en om de Kunera-database met op dit moment vijftienduizend religieuze en profane objecten op lange termijn niet te laten verworden tot een inactieve digitale kaartenbak, is hij recentelijk ontsloten met een Geografisch Informatiesysteem (GIS). Dat was mogelijk dankzij een subsidie van de Stichting Reinier Post. Door het gebruik van een nieuwe GIS-applicatie kunnen de geografische patronen uit het insigne- en ampullenmateriaal op kaarten van Google Maps worden geanalyseerd (www.kunera.nl).

***Brugse Wegwijzer* gevisualiseerd met een Geografisch Informatiesysteem**

Pilots

De nieuwe impuls die zo aan het kunsthistorisch onderzoek kon worden gegeven, kreeg in het najaar van 2008 een vervolg met twee Kleine Data-projecten van DANS: 'Alfa-Geo', mede mogelijk gemaakt door het programma Ruimte voor Geo-informatie. Het doel van die *pilots* was het onderzoek naar middeleeuwse pelgrimsroutes te verbreden door het insignemateriaal binnen de nieuwe internettoepassing samen te brengen met geografische patronen uit heel ander bronmateriaal. Daartoe zijn twee schriftelijke bronnen ontsloten: het *Mirakelboek van Onze Lieve Vrouw van 's-Hertogenbosch* en de Brugse Wegwijzer.

Alle 481 wonderverhalen uit het *Bossche Mirakelboek* (1382-1603), zijn aan de hand van de herkomstplaatsen van de pelgrims op een kaart gezet. Het patroon dat daaruit naar voren komt, is vergeleken met de patronen van Bossche insignes. Die vergelijking bevestigt de toepasbaarheid van geo-technologie in insigneonderzoek, maar ook het nut van de confrontatie met ander historisch bronmateriaal. Het mirakelboek-patroon vult het patroon van de uit 's Hertogenbosch meegenomen insignes aan: de kwetsbare metalen objecten worden namelijk enkel opgegraven in gebieden met een vochtige, luchtafsluitende bodem. Het patroon van de insignes die in 's-Hertogenbosch zijn gevonden, brengt een afwijking aan het licht: het verspreidingspatroon is veel groter dan dat van de eerdergenoemde insignes. De afstand die de reizende Bosschenaar wilde afleggen was klaarblijkelijk groter dan de afstand die de niet-Bosschenaar voor de Bossche Maria wilde overbruggen.

Insigne OLV van 's-Hertogenbosch, vindplaats: Nieuwlande. Bron: R.M. van Heeringen, A.M. Koldeweij en A.A.G. Gaalman, *Heiligen uit de modder. In Zeeland gevonden pelgrimstekens* (Clavis Kunsthistorische Monografieën 4), Utrecht en Zutphen 1987, pp. 78-79, nr. 11.2

Middeleeuwse TomTom

Een nog concretere bron voor middeleeuwse reisroutes is de *Brugse Wegwijzer* (1370-1400). De ontsluiting van deze onderbelichte, maar uiterst waardevolle bron met een Geografisch Informatiesysteem is spectaculair. De routes uit de geschreven bron zijn omgezet in een digitale wegenkaart, waardoor het totaalconcept van deze middeleeuwse routeplanner voor het eerst helder in beeld is gebracht. Daarnaast is de tekst gedigitaliseerd en toegankelijk gemaakt. De geografische patronen van pelgrimssouvenirs met herkomst- en vindplaats Brugge kunnen over het patroon van de wegenkaart worden geplaatst. Vergelijking van beide soorten bronmateriaal heeft ook hier interessante resultaten aan het licht gebracht.

Door de lancering van de internettoepassing Kunera 2.0 is bestaand abstract onderzoeksresultaat zichtbaar gemaakt met behulp van een Geografisch Informatiesysteem. Daardoor is het toegankelijk geworden voor een breder historisch en kunsthistorisch publiek. Het gebruik van geo-technologie is heel waardevol gebleken voor het ontsluiten en vergelijken van historisch bronmateriaal. Kortom, er zit toekomst in het onderzoek naar het verleden.

Op de website www.kunera.nl kunnen de pilot-projecten geraadpleegd worden. Ga via 'direct naar de database' naar het kopje 'Steden in perspectief'.