

---

## E-INFRASTRUCTUREN VOOR CULTUREEL ERFGOED

---

Rosette VANDENBROUCKE

GRID Coordinator, Belnet, Vrije Universiteit Brussel – BEGrid

Elena PHALET

Wetenschappelijke medewerker, Dienst voor Wetenschappelijke en Technische Informatie (DWTI)

- Europese en zelfs wereldwijde e-infrastructuren (netwerken, gedistribueerde computerinfrastructuren, data-infrastructuren) hebben hun opmars gemaakt in het voorbije decennium. Ze openen nieuwe deuren voor onderzoek in alle domeinen. Bandbreedte, rekencapaciteit en opslagcapaciteit die ten dienste staan van onderzoek zijn nooit voordien zo groot geweest. De culturele sector is tot nu toe minder vertrouwd met dit nieuw aanbod maar wil daar verandering in brengen. Internationale projecten in de humane wetenschappen en kunst nemen de uitdagingen van deze nieuwe technologieën aan. In het bijzonder zal het *DC-NET*-project (*Digital Cultural Heritage Network*) nagaan hoe e-infrastructuren op nationaal en internationaal vlak een meerwaarde aan de sector van het culturele erfgoed kunnen bieden en hun onderzoeksprioriteiten stelen op de nieuwe perspectieven die het gebruik van e-infrastructuren opent.
- Les e-Infrastructures européennes et même mondiales (réseaux, infrastructures informatiques distribuées, infrastructures data) ont progressé cette dernière décennie. Elles ouvrent de nouvelles perspectives pour la recherche dans tous les domaines. Bande passante, capacité de calcul et de stockage au service de la recherche n'ont jamais jusqu'ici été si importantes. Le secteur culturel est à ce jour moins familiarisé à cette nouvelle offre, mais va s'y adapter. Des projets internationaux dans les domaines des arts et de sciences humaines relèvent le défi de ces nouvelles technologies. En particulier le projet *DC-NET* (*Digital Cultural Heritage Network*) a pour but de déterminer comment les e-infrastructures nationales et internationales peuvent apporter une valeur ajoutée au secteur du patrimoine culturel et baser leurs priorités de recherche sur les nouvelles perspectives qu'ouvre l'usage des e-infrastructures.

Het *ERA-NET* (*European Research Area Network*) *DC-NET* (*Digital Cultural Heritage Network*)<sup>1</sup>-project, gefinancierd door het 7de Kaderprogramma voor Onderzoek en Ontwikkeling van de Europese Unie (EU), gaat na hoe bestaande e-infrastructuren gebruikt kunnen worden in het domein van digitaal cultureel erfgoed en zal onderzoeksprioriteiten definiëren die gebruik maken van de perspectieven geopend door het gebruik van nationale en internationale e-infrastructuren.

E-infrastructuren zijn echter nog vaak grote onbekenden voor onderzoekers uit de culturele sector. Een eerste opdracht van *DC-NET* was het met elkaar in contact brengen van vertegenwoordigers uit beide sectoren en tevens een aantal voorbeelden te leveren van projecten die al gebruik maken of willen maken van deze technologieën.

De conferentie die op 29 oktober 2010 werd georganiseerd door de Dienst voor Wetenschappelijke en Technische Informatie (DWTI) van het Federaal Wetenschapsbeleid was één van de stappen in deze opdracht. De conferentie, *e-Infrastructures for cultural heritage*<sup>2</sup>, vond plaats in de Koninklijke Bibliotheek van België, in Brussel, een symbolische en architecturaal indrukwekkende locatie voor dit evenement. Onder de aanwezigen, van 27 verschillende nationaliteiten, bevonden zich vertegenwoordigers van ministe-

ries, onderzoeksinstellingen, de culturele sector en de IT-sector. Het publiek maakte kennis met het concept van e-infrastructuren en wat deze kunnen betekenen voor het onderzoek naar en de exploitatie van het cultureel erfgoed. Een aantal projecten illustreerden hun aanzet tot het gebruik of opzetten van e-infrastructuren.

*DC-NET* houdt ook contact met internationale initiatieven die aanbevelingen formuleren in de ontwikkeling van e-infrastructuren (de *e-Infrastructure Reflection Group* of *e-IRG*)<sup>3</sup> en in het gebruik van e-infrastructuren in de sociale en humane wetenschappen en in de kunst (*Coalition of Humanities and Arts Infrastructures and Networks* of *CHAIN*)<sup>4</sup>.

### e-Infrastructuren

E-Infrastructuren kunnen beschreven worden als de term voor technologie en organisaties die gedistribueerde regionale, nationale en internationale samenwerking bevorderen door middel van digitale netwerken, gedistribueerde rekenfaciliteiten en gezamenlijk gebruik van gedistribueerde gegevens. De voornaamste e-infrastructuren zijn digitale netwerken, rekeninfrastructuren en data-infrastructuren.

## Digitale netwerken

Digitale netwerken zijn de basis voor alle andere e-infrastructuren. Zij zorgen voor de verbinding van allerhande digitale apparatuur door middel van bekabelingen, netwerk apparatuur en digitale technologie. Bovenop deze technologie bieden zij een groot aantal diensten aan die dit digitaal netwerk toegankelijk maken voor de gebruiker. In de onderzoekswereld zijn de nationale onderzoeksnetwerken de meest bekende en gebruikte e-infrastructuren.

*Belnet*<sup>6</sup> is het Belgische onderzoeksnetwerk. Het verbindt universiteiten, hogescholen, onderzoekscentra en overheidsdiensten. Via het Belnet-glasvezelnetwerk van meer dan 1500 kilometer over heel België kunnen deze instellingen snel informatie uitwisselen tussen de verbonden instellingen en krijgen bovendien ook een snelle toegang tot Europese en internationale onderzoeksnetwerken en tot internet.

## Rekeninfrastructuren

Bij **grid computing** vormen verschillende computers, die zich geografisch ver van elkaar kunnen bevinden, een virtuele rekenfaciliteit. Gebruikers krijgen toegang tot het geheel of tot een gedeelte van de gezamenlijke reken capaciteit. Worden er ook opslagcapaciteit en bijbehorende diensten voorzien, dan spreken we van een "data/computing" grid. De idee achter grid computing is het feit dat geen enkele computer voor de volle tijd zijn reken capaciteit benut. Door het delen van een rekeninfrastructuur kan de totale beschikbare rekentijd beter benut worden, kunnen meer onderzoekers toegang krijgen tot reken capaciteit en/of kan er meer reken capaciteit gebruikt worden door een onderzoeker die daarvoor een waaier aan nieuwe onderzoeksmogelijkheden ziet open gaan. Het toevoegen van de data component verhoogt de onderzoeksmogelijkheden.

*BEgrid*<sup>6</sup> is de Belgische reken/data grid voor onderzoek. *BEgrid* wordt gecoördineerd door *Belnet*. Universiteiten en onderzoeksinstellingen zijn de leveranciers van de reken- en opslag capaciteit.

*EGI (European Grid Initiative)*<sup>7</sup> coördineert het gridgebeuren op Europees niveau. Daar zijn de nationale grid-infrastructuren, zoals *BEgrid*, de leveranciers van de rekenmiddelen. *EGI* heeft 10.000 gebruikers uit een variëteit van onderzoeksdomeinen, van archeologie tot zoölogie.

Een **supercomputer** is geen e-infrastructuur op zichzelf maar krijgt dit label wanneer supercomputers met elkaar worden verbonden of wanneer de toegang er toe verruimd wordt voor nationaal

of zelfs internationaal gebruik. Een supercomputer is een computer die heel snel kan rekenen, door onder meer efficiënt parallel rekenen op grote schaal (duizenden processoren tegelijk gebruiken) toe te passen.

**Cloud computing** is het laatste nieuwe broertje in de reeks van rekeninfrastructuren. In feite dekt deze benaming verschillende definities. De definitie "*infrastructuur als een dienst*" zal waarschijnlijk het belangrijkste zijn in de onderzoekswereld. Op deze manier wordt een rekeninfrastructuur aangeboden aan de gebruiker als een virtueel platform, waarbij de gebruiker zijn eigen rekenomgeving kan blijven gebruiken. Met rekenomgeving wordt het geheel van beheersystemen, programma's en gegevens bedoeld. Er is nog geen echt aanbod van cloud computing georganiseerd voor de internationale onderzoekswereld.

## Data-infrastructuren

Data-infrastructuren zijn van primordiaal belang voor onderzoek en zeker voor onderzoek in de culturele sector. Data-infrastructuren zijn minder goed gedefinieerd en er bestaat ook geen aanbod van een nationale of een internationale data-infrastructuur over alle disciplines heen. Wel staat vast dat er een groot aantal specifieke data-infrastructuren bestaan, zoals gedigitaliseerde archieven of gedigitaliseerde museumcollecties. Het aan elkaar knopen van deze data-infrastructuren, gecombineerd met een toegang voor een internationaal onderzoekerspubliek kan een weg openen naar volledig nieuw onderzoek.

## De humane wetenschappen en het gebruik van e-infrastructuren

Onderzoekers uit de sociale en menswetenschappen zijn vaak minder bekend met e-infrastructuren maar in de laatste jaren zijn een aantal projecten opgestart die de mogelijkheden van het gebruik van e-infrastructuren, inclusief data-infrastructuren, verkennen. De extra mogelijkheden die deze insteek biedt om data op te slaan en terug te vinden, manipuleren, verrijken, combineren en onderzoeken, maakt dat tevoren onopgemerkte trends en patronen duidelijk worden en nieuwe onderzoeksvragen kunnen gesteld worden.

Dit toenemend gebruik van ICT voor menswetenschappelijk onderzoek is niet enkel een technologisch proces maar evengoed een sociaal en cultureel gebeuren: de visie op onderzoek is volop aan het veranderen, doordat het nu ook plaatsvindt in een sociaal model van uitwisseling

van ideeën en gegevens in plaats van door een-zame onderzoekers in stoffige archieven.

Het Duitse *TextGrid*<sup>8</sup>, een virtuele onderzoeksomgeving voor filologen, linguïsten, musicologen en kunsthistorici is er één van, net als de Europese initiatieven *Council of European Social Sciences Data Archives (CESSDA)*<sup>9</sup>, *Digital Research Infrastructure for the Arts and Humanities (DARIAH)*<sup>10</sup> en *Common Language Resources and Technology Infrastructure (CLARIN)*<sup>11</sup>, die in het kader van een Europese strategie voor onderzoeksinfrastructuren (het *European Strategy Forum on Research Infrastructures of ESFR*)<sup>12</sup> werden opgestart.

*CESSDA* brengt sociaal wetenschappelijke archieven uit 20 Europese landen samen. Op dit moment biedt *CESSDA* ongeveer 25.000 datacollecties aan, van sociologische enquêtes, verkiezingsonderzoek, longitudinale onderzoeksprojecten, opiniepeilingen en volkstellingen. Jaarlijks komen daar nog ongeveer 1000 datacollecties bij. Een 70.000 onderzoekers maken gebruik van deze datasets in hun geheel, of van onderzoeksvragen en variabelen die in de datasets voorkomen.

Het project *DARIAH* start in maart 2011, na 4 jaar voorbereidend werk, met de constructiefase van een e-infrastructuur die het onderzoek in menswetenschappen en kunsten op een ander niveau wil brengen: door middel van hoogtechnologische middelen en methoden kunnen verschillende soorten digitaal materiaal, afkomstig van verschillende plaatsen, worden samengebracht om nieuwe onderzoeksvragen en betere onderzoeksresultaten mogelijk te maken.

*CLARIN* startte in 2008 en wil tegen juni 2011 een prototype bouwen voor een e-infrastructuur voor linguïstisch onderzoek. Er bestaan al een heel aantal technologieën en hulpmiddelen voor dergelijk onderzoek, en *CLARIN's* bijdrage zou erin bestaan deze samen te brengen in een stabiel en duurzaam, interoperabel, voor iedereen toegankelijk en bruikbaar systeem.

In november 2010 werd het project *European Holocaust Research Infrastructure (EHR)*<sup>13</sup> gelanceerd.

Documenten, foto's, objecten en films in verband met de holocaust zijn verspreid over heel de wereld. De komende 4 jaar zullen historici, archivisten en informatici samenwerken om materiaal uit 20 instellingen van 13 verschillende landen via een portaalsite samen te brengen. Onderzoekers zowel als het grote publiek kunnen zo een schat aan informatie van achter hun computer doorzoeken in plaats van grote afstanden te moeten

afleggen om de verschillende archieven zelf te bezoeken.

*Mapping the Republic of Letters*<sup>14</sup> is een project aan de Universiteit van Stanford, dat het verspreiden van ideeën onder geleerden in de periode 1500-1800 grafisch weergeeft, op basis van enorme hoeveelheden correspondentie (Voltaire alleen al schreef meer dan 18.000 brieven), reisverslagen, bibliotheekcatalogi... Door al deze informatie samen te brengen en visueel voor te stellen, komen de onderzoekers tot nieuwe inzichten over hoe die netwerken eigenlijk functioneerden.

## Noden en vereisten van de culturele wereld aangaande data-infrastructuren

Het samenbrengen en opslaan van grote hoeveelheden menswetenschappelijke en culturele data is echter niet altijd een rechtlijnig en voor de hand liggend proces. Deze menswetenschappelijke data werden meestal binnen een bepaalde context verzameld en zijn al van een zekere interpretatie en annotaties voorzien. De psychologische drempel om ze te delen, en er controle over te verliezen, ook al is dat gedeeltelijk, kan groot zijn.

De hoeveelheid informatie groeit constant aan. De culturele wereld heeft dus nood aan data-infrastructuren, die een bewaring op lange termijn garanderen, met de nodige back up-systemen, mogelijkheid tot het correct en recent houden van de informatie ("curation") en met open maar veilige toegang. Deze toegang kan ook verschillend zijn voor verschillende doelgroepen. Onderzoekers krijgen bijvoorbeeld een andere toegang tot de gegevens dan het brede publiek. Een verfijnd en streng systeem voor "identity management" en toegangscontrole tot de data zal dan ook nodig zijn.

Andere vereisten zijn onder meer: het gebruik van standaarden en metadata, krachtige zoekmechanismen voor complexe gegevens, interoperabiliteit, beheer van rechten, gebruik en hergebruik van inhoud, gevorderde diensten voor het beheer van collecties, meertaligheid...

De vooruitgang die nu geboekt wordt in het creëren van culturele data-infrastructuren op internationaal niveau zal wellicht in de nabije toekomst nog meer vereisten aan het licht brengen.

## Overlegorganen

Er bestaan overlegorganen die de problematiek van het gebruik van e-infrastructuren en de creatie van data-infrastructuren behandelen. Twee belangrijke spelers in dit domein zijn *CHAIN* en *e-IRG*.

### *De Coalition of Humanities and Arts Infrastructures and Networks (CHAIN)*

Bovenstaande onderwerpen worden door vertegenwoordigers van e-infrastructuren besproken op bijeenkomsten als die van de *Coalition of Humanities and Arts Infrastructures and Networks, CHAIN*. Dit is een discussieforum zonder eigen budget en met een lichte organisatie structuur: vertegenwoordigers van onderzoeksinfrastructuren voor sociale en menswetenschappen en een lijst van thema's die worden opgevolgd.

Eén van de themata voor discussie is alvast de fragmentarisering van het e-infrastructuurlandschap: onderzoekers werken nog teveel geïsoleerd op hun eigen domein, met onderling incompatibele technologieën. Deze fragmentarisering moet met openheid, communicatie en meer coördinatie worden tegengegaan om de technologische mogelijkheden van e-infrastructuren ten volle te kunnen benutten.

Een andere uitdaging is de duurzaamheid van de bestaande e-infrastructuren: hun financierings- en businessmodellen dus. Ook voor de aanpak van juridische problemen, zoals bijvoorbeeld auteursrechten, is onderling overleg nuttig. En het bereiken en motiveren van gebruikers blijft een belangrijk thema<sup>15</sup>.

### *e-Infrastructures Reflection Group (e-IRG)*

De leden van de *e-IRG* zijn vertegenwoordigers uit de Europese landen, aangeduid door hun respectievelijke minister voor wetenschapsbeleid of equivalent, en buigen zich over de toekomst van e-infrastructuren en aanverwante diensten. *E-IRG* "White papers" behandelen deze onderwerpen en geven aanbevelingen, gericht aan de regeringen, voor hun beleid op korte tot middellange termijn. De *e-IRG* "Roadmap" levert aanbevelingen op lange termijn. Werkgroepen binnen de *e-IRG* rapporteren over specifieke onderwerpen. Interessant te vermelden in deze context zijn het "Data Management Task Force Report" en de "Blue Paper" die geschreven werden ten behoeve van *ESFRI* en aanbevelingen inhoudt voor het gebruik van e-infrastructuren en diensten door *ESFRI*-projecten.

## Het project *Digital Cultural Heritage Network (DC-NET)*

In december 2009 ging het *ERA-NET*-project *Digital Cultural Heritage Network (DC-NET)* van start, gefinancierd door het 7de Kaderprogramma voor onderzoek en ontwikkeling van de Europese Commissie. *DC-NET* heeft als doel te onderzoeken hoe e-infrastructuren kunnen bijdragen tot de ontwikkeling van nieuwe diensten en wil een gemeenschappelijk programma van onderzoeksprioriteiten opstellen voor de digitale culturele erfgoedsector.

De partners van *DC-NET* zijn ministeries voor cultuur en wetenschap van 8 Europese lidstaten, waaronder het Belgische Federaal Wetenschapsbeleid vertegenwoordigd door de Dienst voor Wetenschappelijke en Technische Informatie (DWTI). Het is de bedoeling om dit netwerk uit te breiden door meer instellingen en landen aan te trekken die geïnteresseerd zijn in het samenbrengen van cultureel erfgoed en e-infrastructuren.

Vertegenwoordigers van de culturele erfgoedsector en van e-infrastructuren zullen samen nagaan met welke bestaande en welke extra te ontwikkelen diensten e-infrastructuren een meerwaarde kunnen betekenen voor onderzoek over en voor cultureel erfgoed.

Een eerste taak is uitgebreide informatie te verzamelen over de nationale e-infrastructuren in de betrokken landen, die in aanmerking komen om met de culturele sector samen te werken. Dit zijn onder meer de Nationale netwerken voor onderzoek en onderwijs (NREN's) van de verschillende landen en de computing/data grid-infrastructuren. Aan de hand van de bekomen informatie werd een e-Infrastructuur handboek, het *e-Infrastructures Handbook*<sup>16</sup> opgesteld, een praktische gids voor de culturele erfgoedsector om meer te weten over e-infrastructuren. Deze gids bevat technische informatie<sup>17</sup>, informatie over de kosten van het gebruik van e-infrastructuren en contactinformatie.

*DC-NET* heeft ook aan alle partnerlanden een overzicht gevraagd van de nationale onderzoeksprioriteiten in het onderzoek over cultureel erfgoed, met de bedoeling deze samen te brengen tot een gezamenlijk actieplan voor onderzoek in deze sector.

*DC-NET* heeft na diepte-interviews van experts en een workshop in Parijs een aantal diensten gedefinieerd, die geleverd kunnen worden door e-infrastructuren en waarvoor onderzoekers uit de culturele erfgoedsector vragende partij zijn.

E-infrastructuren kunnen een belangrijke rol spelen in verschillende stadia van het onderzoeksproject: allereerst bij het vinden van informatie, bijvoorbeeld met geavanceerde zoekfuncties, die 2D- en 3D-objecten, audiovisueel materiaal en metadata kunnen localiseren.

Vervolgens zou het analyseren en bewerken van data door gebruik van e-infrastructuren op een ander niveau gebracht kunnen worden. De concrete noden voor analyse en bewerking, zoals bijvoorbeeld indexing, statistische analyse of geografische localisatie, zijn evenwel niet dezelfde voor verschillende wetenschappelijke onderzoeken. Hierover moet nog verder een verdere bevraging gebeuren.

Ten derde kunnen e-infrastructuren in het ter beschikking stellen van informatie en in het uitwisselen over onderzoeksresultaten een rol spelen. Er bestaan veel technologische tools die samenwerking vergemakkelijken en die door e-infrastructuren als diensten kunnen worden aangeboden, zoals tele- en videoconferenties. De culturele erfgoedsector is ook vragende partij voor een technologische omgeving waarin enorme hoeveelheden digitaal materiaal veilig en duurzaam kunnen worden opgeslagen, met de nodige back ups en identificatie-maatregelen om gebruikers correct te identificeren. De bijkomende diensten die de culturele erfgoedsector van e-infrastructuurbeheerders zou-

den vragen om aan deze noden tegemoet te komen, zullen vervolgens op hun technische, financiële en praktische haalbaarheid beoordeeld worden.

Tegen eind november 2011, het einde van het *DC-NET* project, zal er belangrijk werk geleverd zijn om een e-infrastructuur voor cultureel erfgoed te definiëren die nieuwe deuren opent voor het onderzoek en de exploitatie van het digitaal cultureel erfgoed.

#### **Rosette Vandenbroucke**

*Belnet – BEGrid*  
Platinum Gebouw  
Louizalaan, 231  
1050 Brussel  
Rosette.Vandenbroucke@belnet.be

#### **Elena Phalet**

*Dienst voor Wetenschappelijke en  
Technische Informatie (DWTI)*  
Platinum Gebouw  
Louizalaan, 231  
1050 Brussel  
Elena.Phalet@stis.belspo.be

*Februari 2011*

## Noten

- 1 Het programma en alle presentaties zijn terug te vinden op:  
*Digital Cultural Heritage Network* [online]. <<http://www.dc-net.org>> (geraadpleegd op 27 februari 2011).
- 2 Digital Cultural Heritage Network. *E-Infrastructure for Cultural Heritage* [online]. <<http://www.dc-net.be>> (geraadpleegd op 27 februari 2011).
- 3 *e-Infrastructure Reflexion Group* [online]. <<http://www.e-irg.eu>> (geraadpleegd op 27 februari 2011).
- 4 *CHAIN, Co-ordination & Harmonisation of Advanced e-Infrastructure* [online]. <<http://www.chain-project.eu>> (geraadpleegd op 27 februari 2011).
- 5 *Belnet* [online]. <<http://www.belnet.be>> (geraadpleegd op 27 februari 2011).
- 6 *BEgrid, The Belgian grid for Research* [online]. <<http://www.begrid.be/>> (geraadpleegd op 27 februari 2011).
- 7 *European Grid Infrastructure* [online]. <<http://www.egi.eu>> (geraadpleegd op 27 februari 2011).
- 8 Zie de presentatie van Heike Neuroth op de DC-NET conferentie: Neuroth, H. *TextGrid – A virtual research environment for humanities* [online], In DC-NET. *e-Infrastructures for cultural heritage*. Brussel, 29 oktober 2010 (geraadpleegd op 27 februari 2011) <<http://www.dc-net.be/en/Programme/heike-neuroth.html>>
- 9 *CESSDA, Council of European Social Science Data Archives* [online]. <<http://www.cessda.org/>> (geraadpleegd op 27 februari 2011).
- 10 *DARIAH, Digital Research Infrastructure for the Arts and Humanities* [online]. <<http://www.dariah.eu/>> (geraadpleegd op 27 februari 2011).

- 11 *CLARIN, Common Language Resources and Technology Infrastructure* [online]. <<http://www.clarin.eu/external/>> (geraadpleegd op 27 februari 2011).
- 12 European Commission. *ESFRI* [online]. <[http://ec.europa.eu/research/infrastructures/index\\_en.cfm?pg=esfri](http://ec.europa.eu/research/infrastructures/index_en.cfm?pg=esfri)> (geraadpleegd op 27 februari 2011).
- 13 EHRI, European Holocaust Research Infrastructure [online]. <<http://www.ehri-project.eu/>> (geraadpleegd op 27 februari 2011).
- 14 Stanford University. *Mapping the Republic of Letters* [online]. <<https://republicofletters.stanford.edu/>> (geraadpleegd op 27 februari 2011).
- 15 Zie het panelgesprek op de "Digital Humanities 2010"- conferentie: Wynne, M. *CHAIN Panel Session at DH 2010* [online]. <[http://www.dariah.eu/index.php?option=com\\_wordpress&p=77&Itemid=198](http://www.dariah.eu/index.php?option=com_wordpress&p=77&Itemid=198)> (geraadpleegd op 27 februari 2011).
- 16 Binnenkort beschikbaar op <<http://www.dc-net.org>>.
- 17 Dat wil zeggen. technische gegevens over netwerkinfrastructuren, data-infrastructuren en computing infrastructu-  
ren, en informatie over hoe men er in praktijk toegang toe krijgt.