

Digital Humanities en de UBVU: vooronderzoek

Michiel Cock

<i>versie</i>	<i>datum</i>	<i>wijzigingen</i>
1.0	18-6-2015	eerste versie voor bespreking Arjan Schalken
1.1	29-7-2015	wijzigingen na eerste bespreking Arjan Schalken doorgevoerd
1.2	11-9-2015	suggesties Ana verwerkt; tweede & derde ronde suggesties Arjan verwerkt; suggesties Sebastien verwerkt
1.3	18-9-2015	suggesties Wouter Gerritsma verwerkt
1.4	21-9-2015	suggestie Josje Calff verwerkt: onderverdeling in prioritering binnen roadmap

Managementsamenvatting

Digital humanities is voor de UB als contentmanager een belangrijke ontwikkeling binnen de geesteswetenschappen. Omdat de mogelijkheden niet uitgekristalliseerd en helder zijn, is de behoefte geconstateerd aan een verkenning van de mogelijkheden voor dienstverlening op het gebied van digital humanities.

Digital humanities is niet geheel nieuw voor de UB: ondersteuning bij GIS en geodata spelen en de bijdrage aan het project INVENiT² zijn vormen van dienstverlening op het gebied van digital humanities.

In de verkenning zijn op basis van gesprekken met digital humanists bij de VU en literatuuronderzoek naar ervaringen bij andere bibliotheken een aantal kansrijke onderwerpen naar boven gekomen. Op basis van een enquête onder geïnteresseerden in digital humanities kom ik tot het volgende voorstel voor prioritering en aanpak:

1. data
 1. digitalisering op verzoek: initiatieven nemen om projecten op te starten
 - a. advies, bemiddeling bij verwerven van externe (secundaire) data: initiatieven nemen om projecten op te starten
 - b. beschikbaar stellen daarvan: activiteiten van anderen volgen, waar kansen zijn onze expertise inbrengen
 3. advies over (meta)data: expertise verder ontwikkelen
2. infrastructuur
 1. verbeteren Beeldbank: initiatieven nemen om projecten op te starten
 2. platform om data in op te slaan, te analyseren en te bewerken: activiteiten van anderen volgen, waar kansen zijn onze expertise inbrengen
 3. repository voor tools (of andere resultaten van onderzoek naast data en publicaties): actief geconstateerde behoefte onder de aandacht brengen bij andere stakeholders en in samenwerking uitwerken op welke manier aan de behoefte voldaan kan worden
3. informatievoorziening, onderwijs & training
 1. tools en technieken: actief geconstateerde behoefte onder de aandacht brengen bij andere stakeholders en in samenwerking uitwerken op welke manier aan de behoefte voldaan kan worden
 2. verbinden van onderzoekers: actief geconstateerde behoefte onder de aandacht brengen bij andere stakeholders en in samenwerking uitwerken op welke manier aan de behoefte voldaan kan worden
 3. crowdsourcing: niet actief vervolgen, maar op verzoek ondersteunen en opschalen waar nodig

Deze prioritering en aanpak is verwerkt in een gefaseerde roadmap met activiteiten en projecten, te vinden op p. 14-16.

1. Inleiding

Digital humanities is voor de UB een belangrijke ontwikkeling binnen de geesteswetenschappen. Omdat de mogelijkheden voor verandering nog niet uitgekristalliseerd en helder zijn, is de behoefte geconstateerd aan een verkenning van de mogelijkheden voor dienstverlening op het gebied van digital humanities. Ik heb deze verkenning uitgevoerd en daarvoor de volgende werkwijze gekozen:

- een vooronderzoek naar mogelijke vormen van dienstverlening
- voorstel aan het MT met een aantal kansrijke onderwerpen
- uitwerken van goedgekeurde onderwerpen in deelprojecten

Het vooronderzoek is gedaan op basis van literatuur en verkennende gesprekken met onderzoekers. Dit document is daarvan de weerslag en bevat:

- een globale business case voor dienstverlening op het gebied van digital humanities
- een inkadering van de verkenning
- kansrijke onderwerpen, inclusief een onderbouwing voor dienstverlening op dat gebied
- prioritering van kansrijke onderwerpen, inclusief een voorstel voor een roadmap

Van het MT wordt akkoord gevraagd met prioritering van de kansrijke onderwerpen, op basis waarvan deelprojecten kunnen worden uitgewerkt. Daarin kan nieuwe dienstverlening ontwikkeld worden of, wanneer het gaat om een pilot, aanvullende informatie voor een business case verzameld worden.

2. Globale business case

In het beleidsplan 'Hotspot voor content. Meerjarenbeleidsplan 2014-2016' van de Universiteitsbibliotheek wordt op meerdere punten het onderwerp e-humanities genoemd.¹ Het is één van de actuele, brede thema's (naast bijvoorbeeld datamanagement, publicatieadvies, geografische informatiesystemen, digitale leer- en onderzoeksomgevingen, multimediale technieken, weblectures en massive open online courses (MooC's)), waarvan kennis nodig is bij contentspecialisten.

Digital humanities is een ontwikkeling binnen de geesteswetenschappen die van belang is voor de UB, omdat het de omgang met content in veel onderdelen van het onderzoeksproces beïnvloedt. Onderzoekers maken op een andere manier gebruik van andere content² en creëren andere output³. Omdat het in het beleidsplan de ambitie van de UB is om de gehele informatieketen te ondersteunen, moet de dienstverlening van de bibliotheek meebewegen met veranderingen in het primaire proces van de onderzoekers. Hoe die dienstverlening op het gebied van digital humanities moet veranderen is nu nog niet duidelijk.

Zoals bij veel bibliotheken is het ook bij de UBVU wel duidelijk dat digital humanities een ontwikkeling is met impact voor de bibliotheek. Tegelijkertijd zijn de mogelijkheden voor verandering nog niet uitgekristalliseerd en helder. Een bijeenkomst voor iedereen in Nederland en

¹pp. 6, 20 en 23.

²"Humanities research has evolved and requires more sophisticated and scalable access to digital sources, such as batch-downloading a digitized corpus or text-mining millions of digital books." Schaffner, Jennifer, and Ricky Erway. 2014. Does Every Research Library Need a Digital Humanities Center? Dublin, Ohio: OCLC Research (<http://www.oclc.org/content/dam/research/publications/library/2014/oclcresearch-digital-humanities-center-2014.pdf>).

³Om een nieuwe visie te ontwikkelen op wat een 'scholarly publication' is, is bijv. één van de aanbevelingen in het rapport van Williford et al., One Culture. Computationally Intensive Research in the Humanities and Social Sciences. A Report on the Experiences of First Respondents to the Digging Into Data Challenge, CLIR 2012.

Vlaanderen die zich bij een onderzoeksbibliotheek (zoals in Nederland UB's of de KB) betrokken is of zou willen worden bij digital humanities bevestigde dat beeld: alle aanwezigen beaamden dat het een belangrijke ontwikkeling is, maar nog weinigen hadden dat omgezet in concrete diensten.

Digital humanities is niet nieuw voor de UB. GIS en geodata spelen ook vaak een rol in digital humanities projecten. De ondersteuning die de UB in de vorm van Geoplaza hiervoor biedt, is ook al een vorm van dienstverlening op het gebied van digital humanities. Sinds januari 2015 is de UBVU ook partner in het project INVENIT²⁴. De faculteiten Geesteswetenschappen en Computer Science hebben dit geïnitieerd. Het project ambieert een betere ontsluiting van online erfgoedcollecties. Door (bestaande) zoekfunctionaliteiten te verbeteren en met elkaar te verbinden kunnen onderzoekers deze bronnen bevragen op manieren die voorheen onmogelijk waren. De UBVU, in het bezit van veel en veelal uniek religieus erfgoed, zorgt met ca. 300 gedigitaliseerde bijbelillustraties voor content in dit project. Ook met de metadata bij deze illustraties levert zij een belangrijks grondstof. Tot slot werft de UBVU mogelijke 'crowdsourcers', o.a. uit de Stichting Vrienden UBVU, en begeleidt hen bij een 'crowdsourcingspilot' binnen INVENIT². Het project loopt tot najaar 2015.

Het ambitieniveau van de VU op het gebied van digital humanities is hoog. In samenwerking met de UvA, KNAW en IBM wordt gewerkt aan het instellen van een Center for Humanities and Technology (CHAT)⁵. Digital humanities is een belangrijk onderdeel van het profileringsthema 'Connected World' uit het Instellingsplan 2015-2020. Het wordt ook genoemd als één van de speerpunten van onderwijs en onderzoek bij FGW⁶. Om bij te dragen aan deze innovatie in onderwijs en onderzoek moet daarbij ook het ambitieniveau van de UBVU hoog zijn. Op dit moment wordt vooral een investering in tijd gevraagd.

In het meerjarenbeleidsplan worden vijf hoofdlijnen van producten en diensten onderscheiden:⁷

1. Bibliotheeklocaties: UB en MB
2. Wetenschappelijke content als grondstof: collecties en toegang
3. Informatievaardigheden: omgang met data en informatie in een academische omgeving
4. Gespecialiseerde contentgerelateerde diensten
5. Infrastructuur voor en managementinformatie over content

Digital humanities wordt onder punt 4 genoemd als één van de gespecialiseerde contentgerelateerde diensten. Tegelijkertijd bieden de andere hoofdlijnen ook richting voor producten en diensten op het gebied van digital humanities. Binnen de VU is voor digital humanities niet voor al die hoofdlijnen op dit moment een concrete vraag geconstateerd. Aan een fysieke locatie (bijvoorbeeld in de vorm van een Digital Humanities Center⁸) lijkt geen dringende behoefte te bestaan. In het vervolg worden kansrijke onderwerpen van dienstverlening onderverdeeld in drie categorieën, in overeenkomst met de drie overige hoofdlijnen:

1. data
2. infrastructuur
3. informatievoorziening, training & onderwijs

⁴Meer informatie op <http://invenit.wmprojects.nl/>.

⁵<https://www.knaw.nl/shared/resources/actueel/publicaties/pdf/meaning-and-perspectives-in-the-digital-humanities-white-paper-chat>.

⁶Bijvoorbeeld <http://fgw.vu.nl/nl/opleidingen/graduate-school-of-humanities/index.asp> en <http://fgw.vu.nl/nl/opleidingen/graduate-school-of-humanities/samenwerking/index.asp>.

⁷p. 16.

⁸Bij andere universiteiten en bibliotheken voldoet dit wel aan een behoefte, bijvoorbeeld University of Virginia Library Scholars' Lab, <http://scholarslab.org/>.

3. Definitie

3.1 Doelgroep

De innovatie van 'digital humanities' is inmiddels zo verspreid dat het niet meer alleen de 'innovators', maar ook 'early adopters' betreft. Dat betekent dat ook de potentiële doelgroep voor diensten groeit. Het is bij het bepalen van beleid óók belangrijk om de doelgroep te segmenteren en rekening te houden met het onderscheid tussen 'innovators', 'early adopters' en op termijn de 'early majority'. Bijvoorbeeld op het gebied van kennisdeling verschillen de behoeften van deze groepen.

De innovators, degenen die zich nu al bezig houden met digital humanities, zijn makkelijk te identificeren: dit zijn bijvoorbeeld de mensen die onderzoek presenteren op congressen als DHBenelux. De groep 'early adopters', degenen die op dit moment het idee hebben dat ze 'iets' met digital humanities willen maar nog niet weten wat dan precies, is lastiger in kaart te brengen. Daarvoor is goede kennis van het onderzoeksveld, op basis van het netwerk van vak/themaspecialisten, onontbeerlijk.

Digital humanists van de VU maken deel uit van verschillende organisatieonderdelen. Het merendeel werkt bij FGW, FGG en FEW. Daarnaast zijn ook onderzoekers van FSW betrokken. Binnen de faculteiten zijn de 'innovators' vooral te vinden in de volgende organisatie-onderdelen:

- Computational Lexicology & Terminology Lab (FGW, Piek Vossen)
- Eep Talstra Centre for Bible and Computer (FGG, Wido van Peursen)
- Geschiedenis (FGW, Inger Leemans, Susan Legêne, Fred van Lieburg)
- Mediastudies (FSW, Wouter van Atteveldt)
- Business Web & Media (Computer Science FEW, Lora Aroyo)

Een interdisciplinair organisatie-onderdeel is het Network Institute. Daarnaast participeert de VU (samen met UvA en KNAW) in het Centre for Digital Humanities.

3.2 Uitsluitingen

Buiten het bereik van deze verkenning vallen:

- GIS, omdat dienstverlening op dat gebied al gebeurt door middel van Geoplaza
- ondersteuning van databeheer en archivering van onderzoeksgegevens, omdat dat onderdeel is van het project Datamanagement

Deze onderwerpen vallen buiten de eigenlijke verkenning, maar uiteraard zullen gesignaleerde kansen voor de UB op deze gebieden intern besproken worden.

3.3 Relaties

Voor technische ondersteuning zijn we afhankelijk van UCIT, voor zover die niet door de afdeling DD&I of externe samenwerkingspartners geleverd kan worden. Gezien de samenwerking tussen VU en UvA op het gebied van digital humanities ligt ook samenwerking van de UBvU en UBA bij de ondersteuning voor de hand.

Voor infrastructuur zijn er grote, algemene projecten zoals bijvoorbeeld CLARIAH. UKB wordt genoemd als een van de partijen die dit project ondersteunt, maar onduidelijk is wat deze ondersteuning precies behelst.

In het project Datamanagement wordt generieke ondersteuning op het gebied van onderzoeksdata ontwikkeld.

Ondersteuning op het gebied van digital humanities wordt in minimale mate ook geboden door het Bureau Informatisering van FGW (onder leiding van Eric Akkerman). Zij bieden bijvoorbeeld ondersteuning bij het opzetten van databases en onderhouden een werkbank met achtergrondinformatie over digital humanities⁹.

Alle dienstverlening op dit gebied wordt ontwikkeld in samenwerking met onderzoekers. Waar mogelijk is dat direct in aansluiting bij een specifiek onderzoeksproject; vervolgens kan op basis daarvan meer generieke ondersteuning ontwikkeld worden.

4. Kansrijke onderwerpen

4.1 Beschrijving

In het vooronderzoek zijn op basis van gesprekken met digital humanists bij de VU en literatuuronderzoek naar ervaringen bij andere bibliotheken de volgende kansrijke onderwerpen naar boven gekomen.

1. data
 1. digitalisering op verzoek
 2. advies, bemiddeling bij verwerven van externe (secundaire) data, beschikbaar stellen daarvan
 3. advies over (meta)data
2. infrastructuur
 1. verbeteren Beeldbank
 2. platform om data in op te slaan, te analyseren en te bewerken
 3. repository voor tools (of andere resultaten van onderzoek naast data en publicaties)
3. informatievoorziening, onderwijs & training
 1. tools en technieken
 2. verbinden van onderzoekers
 3. crowdsourcing

4.2 Onderbouwing

4.2.1 Data

Het aanbieden van dataresources is een kernactiviteit van de UB. Daarom zijn deze onderwerpen een logische uitbreiding van onze huidige activiteiten. Dat betekent niet dat alle kennis voor alle onderwerpen al in huis is: bijvoorbeeld op het gebied van Linked Open Data moet onze expertise nog groeien.

4.2.1.1 Digitalisering op verzoek

Noodzakelijke voorwaarde voor het doen van digitaal onderzoek is dat het onderzoeksmateriaal ook digitaal beschikbaar is. Het digitaliseren van algemeen materiaal ligt niet voor de hand: veel meer partijen zijn in staat om dat te digitaliseren. Voor het bijzondere materiaal dat bij de UB/VU aanwezig is daarentegen zijn wij de enige, of een van de weinige, die de beschikking hebben over dat materiaal en daarom moeten we ons bij digitalisering daarop richten.

Voor digitalisering is geen structureel budget beschikbaar. Het kan daarom alleen vraaggericht en met financiering door de aanvrager (of diens subsidiegever) gebeuren. Die onderzoekers die gebruik willen maken van gedigitaliseerde Bijzondere Collecties zijn niet noodzakelijkerwijs verbonden aan

⁹http://www2.let.vu.nl/werkbanken/literatuurcultuur/themas/ehumanities/introductie_.php

de VU. Het is daarom belangrijk dat er een goed overzicht is van wat er in onze Bijzondere Collecties aanwezig is en welk materiaal interessant kan zijn voor deze vormen van onderzoek, om op die manier het gebruik van Bijzondere Collecties in onderzoek te stimuleren.

4.2.1.2 Advies, bemiddeling bij verwerven van externe data, beschikbaar stellen van data

Voor digital humanities-onderzoek wordt ook gebruik gemaakt van data van externe, soms commerciële aanbieders. Piek Vossen bijvoorbeeld maakt gebruik van content van LexisNexis. Hij heeft hiervoor zelf de rechten in orde gemaakt. De UB heeft óók een abonnement op dit bestand; het is daarom logisch om de verdergaande behoefte van digital humanities-onderzoekers mee te nemen in de licentieonderhandelingen. Bovendien worden er nu datasets aangeschaft (bijvoorbeeld door de groep van Piek Vossen) waarvan de licentie het gebruik door anderen binnen de VU toe zou staan. Omdat de dataset voor anderen niet vindbaar en toegankelijk is, wordt de dataset niet zo gebruikt als kan.

Ook bij het verwerven van licenties om contentmining toe te passen op externe bestanden kan de bibliotheek een rol spelen. Contentmining is een ontwikkeling die in alle wetenschapsgebieden in opkomt is; ondersteuning daarbij is daarom niet alleen relevant voor digital humanists. Voor onderzoekers is de complexiteit rondom licenties voor contentmining ontmoedigend. LIBER lobbyt voor aanpassingen aan copyright, zodat dergelijke onderhandelingen over licenties niet meer nodig zijn, onder meer door middel van de The Hague Declaration¹⁰ en het beschikbaar maken van een fact sheet¹¹. UKB heeft in juli deze The Hague Declaration ondertekend. De UB kan hierbij een zelfde rol op zich nemen als bij Open Access tot publicaties, door aandacht te vragen voor het onderwerp en (voortdurend) uit te leggen wat er nodig is en waarom. Vergelijkbaar met Open Access is het dan een taak van de UB om bij elke licentieonderhandeling het thema contentmining ter tafel te brengen en te pleiten voor het recht daarop.

Voor sommige vormen van data-onderzoek is meer rekenkracht nodig dan wat een gewone computer kan bieden. Niet iedereen die deze behoefte heeft, weet hoe die te realiseren. Research Data Services heeft bijvoorbeeld geconstateerd dat voor MA-studenten van FEWEB aanvullende ondersteuning nodig is. Ook voor digital humanists kan advies of bemiddeling door de UB op dit gebied nuttig zijn.

4.2.1.3 Advies over (meta)data

Vanuit de klassieke bibliotheekexpertise op het gebied van catalogiseren kan de UB zich ontwikkelen als specialist op het gebied van metadata en metadata-standaarden.¹² Op die manier kunnen we bijdragen aan de functionaliteit tijdens het onderzoeksproject, maar ook aan interoperabiliteit (bijvoorbeeld door Linked Open Data) en langdurige beschikbaarheid na afloop van het project. In het project INVENiT² worden nieuwe metadata geproduceerd bij de gedigitaliseerde bijbelillustraties. Na afloop van het project zullen deze metadata ook toegevoegd worden aan de Beeldbank. Daarvoor is het belangrijk om in een vroeg stadium na te denken over standaarden en technisch formaat, om het importeren in het systeem van de Beeldbank zo eenvoudig mogelijk te maken.

4.2.2 Infrastructuur

Omdat het hier gaat om digitale infrastructuur, is samenwerking met UCIT (of externe partners) en de daarbij behorende afspraken over rolverdeling essentieel. Technisch, applicatie- en systeembeheer hoort daarbij eerder bij anderen dan bij de UB. Voor de UB is vooral een rol weggelegd in ondersteuning, innovatie en ontwikkeling, contentbeheer en functioneel beheer.

¹⁰<http://thehaguedeclaration.com/>.

¹¹<http://libereurope.eu/wp-content/uploads/2014/11/Liber-TDM-Factsheet-v2.pdf>.

¹²Hotspot, p. 9.

4.2.2.1 Verbeteren Beeldbank

Materiaal dat reeds gedigitaliseerd is, wordt nu aangeboden via de Beeldbank. Voor digital humanities-onderzoek biedt de interface daarvan niet genoeg mogelijkheden. Onderzoekers willen bijvoorbeeld over de tekstbestanden zelf kunnen beschikken om daarop hun eigen software los te laten. Een concrete vraag hiervoor is al binnengekomen: Pim Huijnen (UU) wilde graag de beschikking hebben over alle bestanden van het tijdschrift 'De accountant' voor historisch onderzoek naar de invloed van de Verenigde Staten op het Nederlandse denken over bedrijfsvoering in de twintigste eeuw. Dit maakt ook duidelijk dat gedigitaliseerd aanbod niet alleen voor de VU, maar ook voor onderzoekers daarbuiten interessant is. Het laat ook zien dat het materiaal interessant is voor anderen dan men op het eerste gezicht verwacht: in dit geval wordt een economisch tijdschrift gebruikt voor historisch onderzoek.

4.2.2.2 Platform om data in op te slaan, te analyseren en te bewerken

Veel digital humanists zijn in staat om hun eigen digitale gereedschap te ontwikkelen of hebben (tijdelijk) de beschikking over professionele ontwikkelaars. Naarmate methodes van digital humanities meer ingeburgerd raken, groeit de groep van onderzoekers die daartoe niet in staat is. Voor hen zou een platform met data (en/of een simpele manier om eigen data toe te voegen) en standaardgereedschap een uitkomst zijn. Er zijn ook vergelijkbare initiatieven elders, bijvoorbeeld DHBox¹³, waarbij een standaardconfiguratie van tools wordt aangeboden. Voor het werken met geodata is Geoplaza een vergelijkbaar platform dat mede door de UB onderhouden wordt. Tegelijkertijd worden er door onderzoekers tools ontwikkeld die in sommige gevallen een dusdanig groot publiek weten te bereiken, dat het ondersteunen van het gebruik een taak is die onevenredige druk legt op de inzet van die onderzoekers. Voor Natural Language Processing is die behoefte geconstateerd door Piek Vossen en Wouter van Atteveldt; zij hebben daarom een eerste projectvoorstel geschreven voor een NLP e-lab, dat besproken is in de Commissie IT & Research. Onderzocht moet worden welke onderzoeksactiviteiten een platform moet bedienen en of daarvoor een eigen infrastructuur nodig is.

4.2.2.3 Repository voor tools (of andere resultaten van onderzoek naast data en publicaties)

Naast meer traditionele resultaten als tijdschriftartikelen of boeken worden in digital humanities-projecten vaak ook andere resultaten opgeleverd, zoals bijvoorbeeld een speciaal voor het onderzoek ontwikkelde tool of een demonstrator voor het gebruik van data. In het veld worden die tools meer en meer langs dezelfde meetlat gelegd als traditionele vormen van onderzoeksresultaten, zoals blijkt uit de vraag om beter ontwikkelde vormen van 'tool criticism' (het kritisch beoordelen van gemaakte tools), zoals er voor boeken een systeem van recensies bestaat. Ook in het rapport 'Does Every Research Library Need a Digital Humanities Center?' wordt deze behoefte gesignaleerd¹⁴. De beschikbaarheid van deze nieuwe vorm van resultaten is op langere termijn niet gegarandeerd, omdat ze op dit moment vaak alleen bij de individuele onderzoeker bewaard worden. Het kan een logische uitbreiding zijn van de rol die bibliotheken hebben bij het beschikbaar houden van wetenschappelijke output (traditioneel) in de vorm van boeken en tijdschriften en (meer recent) van onderzoeksdata. Tools kunnen bijvoorbeeld in de vorm van een Virtual Machine in een repository opgeslagen worden. Daarmee is de code inclusief alle afhankelijkheden (van bijvoorbeeld een specifiek besturingssysteem of andere software) voor belangstellenden beschikbaar; de onderhoudslast voor de bibliotheek wordt daarmee sterk beperkt. Gebruikers kunnen een tool uit de

¹³<https://github.com/DH-Box/dhbox/wiki>

¹⁴"Many academics would like to include their untraditional scholarly outputs in online catalogs so that they will become discoverable alongside other humanities resources." Schaffner, Jennifer, and Ricky Erway. 2014. Does Every Research Library Need a Digital Humanities Center? Dublin, Ohio: OCLC Research (<http://www.oclc.org/content/dam/research/publications/library/2014/oclcresearch-digital-humanities-center-2014.pdf>).

repository downloaden en op hun eigen computer in een Virtual Machine gebruiken. Drie andere ontwikkelingen zijn relevant bij dit onderwerp. Veel ontwikkelaars maken broncode beschikbaar via platforms als GitHub of Bitbucket. Dat is echter nog niet hetzelfde als een werkende tool beschikbaar stellen. Als tweede moet vermeld worden dat bij het archiveren van data in de VU data repository ook software bijgevoegd mag worden wanneer die noodzakelijk is om de databestanden te kunnen openen. Als laatste is deze behoefte ook door anderen geconstateerd: bijvoorbeeld bij DANS is aandacht voor dit onderwerp.

4.2.3 Informatievoorziening, onderwijs & training

4.2.3.1 Tools en technieken

De UB heeft een rol bij het leren omgaan met content. Informatievaardighedenonderwijs is één van de kerntaken van de UB. Wanneer onderzoek een andere manier van werken met content vraagt, zoals bij digital humanities, ligt het in de lijn der verwachting dat ook ons onderwijs daarop wordt aangepast. Zoals het informatievaardighedenonderwijs bij faculteiten vaak is ingebed in het bredere kader van academische vaardigheden, kan ook training in tools en technieken op het gebied van digital humanities breder ingebed zijn in het onderwijs van de faculteiten. Daarnaast kan de UB kennis op het gebied van informatievaardigheden inzetten om de opkomende ontwikkeling van 'tool criticism' vorm te geven.

Het is daarnaast ook mogelijk om als UB de rol van innovatieve partner op ons te nemen, door training te organiseren waarvoor belangstelling bestaat, maar waarvoor (nog) geen plek is in het formele curriculum van studenten. Te denken valt aan een training in Python die zich specifiek richt op geesteswetenschappers of training in data-visualisatie. In ieder geval binnen de groep van Inger Leemans bestaat de behoefte aan een training Python. De UB heeft niet de kennis in huis om dergelijk onderwijs geheel zelf te verzorgen; in plaats daarvan kunnen wij de organisatie op ons nemen, door bijvoorbeeld geschikte docenten te vinden en een locatie beschikbaar te stellen. Voor sommige onderwerpen kunnen we op basis van onze eigen ervaring presentaties ontwikkelen, bijvoorbeeld over digitalisering van bronnen op basis van onze ervaringen met digitaliseringsprojecten.

Voor informatievoorziening op het gebied van digital humanities in het algemeen is er de werkbank van Bureau Informatisering van FGW. Daarin staat echter geen specifieke informatie over tools en technieken, zoals voor welke soort vraag welke tool het meest gebruikt of geschikt is. De UB kan bijvoorbeeld in de vorm van een Libguide een dergelijk overzicht samenstellen.

Literatuuronderzoek door de bibliotheek is bij FGW geen gebruikelijke dienst, omdat onderzoekers binnen hun eigen vakgebied voldoende informatievaardig zijn. Bij een nieuw gebied als digital humanities hebben early adopters wél behoefte aan die vorm van ondersteuning. In gesprekken met Marije Martijn en Henk de Regt (FGW, afdeling Wijsbegeerte) gaven zij daarvan voorbeelden aan de hand van concrete vragen over de toepasbaarheid van technieken uit Digital Humanities voor hun specifieke onderzoeksvragen (bijv. 'Kan topic modeling toegepast worden op Arabische teksten?').

4.2.3.2 Verbinden van onderzoekers

Onderzoekers zijn vaak uitstekend in staat om elkaar te vinden op basis van inhoudelijke raakvlakken: onderzoekers van retorica in de Oudheid leggen makkelijk contact met degenen die retorica in andere tijdvakken onderzoeken. Wanneer er technische of methodologische raakvlakken zijn, worden die verbindingen veel minder snel gelegd. Digital.Humanities@Oxford (waarin ook de Bodleian Library deelneemt) onderhoudt bijvoorbeeld een overzicht van projecten waarbij een gebruiker kan zoeken op 'ICT Method'. Op die manier kunnen onderzoekers zelf gemakkelijk technische expertise van anderen vinden. Als beheerder van het CRIS en de repository speelt de UB al een rol in het beheer van de gegevens op basis waarvan de expertise van onderzoekers vastgesteld kan worden. Speciaal binnen digital humanities kan hiervoor extra dienstverlening ontwikkeld worden, bijvoorbeeld door in samenwerking met het Centre for Digital Humanities te

werken aan een overzicht van projecten & expertise. In de gesprekken met Marije Martijn en Henk de Regt kwam naar voren dat zij als early adopters behoefte hebben aan een laagdrempelige gesprekspartner op het gebied van Digital Humanities. Een dergelijk overzicht kan helpen om zo'n gesprekspartner te vinden.

Doordat vak/themaspecialisten met veel verschillende onderzoekers en docenten spreken, kunnen wij verbindingen leggen die anderen niet altijd zien. Een recent voorbeeld is bijvoorbeeld het maken van digitale tekstedities: het blijkt dat bij verschillende vakken van verschillende docenten (Nelleke Moser, August den Hollander, Ton Derks) studenten wordt gevraagd een digitale teksteditie te maken, zonder dat de docenten dat van elkaar weten. De UB kan een rol spelen in kennisuitwisseling door bijeenkomsten over die onderwerpen te organiseren, waarvoor we bij meerdere onderzoekers of docenten belangstelling hebben geconstateerd.

Als laatste kan de UB ook deze rol spelen door voor onderzoekers de gelegenheid te organiseren om projecten of projectresultaten aan anderen te presenteren.

4.2.3.3 Crowdsourcing

In de loop van de jaren heeft de UBvU een achterban verzameld die affiniteit heeft met en kennis van de traditie waaruit de universiteit is voortgekomen, waaronder de Vrienden UBvU. Met deze expertise vormt zij (een mogelijke) 'crowd' die de metadata van religieus erfgoed kan verrijken. Deze optie wordt, via een pilot, verkend in het project INVENIT², waaraan ook de UBvU deelneemt. Eén van de projectdoelen is om emoties in religieus erfgoed uit de vroegmoderne tijd inzichtelijk te maken. De verwachting is dat de 'crowd', met haar vertrouwdheid met de protestantse traditie, een dataset kan creëren waarmee historici aan de slag kunnen. Deze werkhypothese wordt getetst in het project. Mogelijk biedt de pilot aanknopingspunten om deze werkwijze toe te passen in toekomstig geesteswetenschappelijk onderzoek.

Tabel: mogelijke dienst per groep uit de innovatie life cycle

	<i>innovators</i>	<i>early adopters</i>
1.1. digitalisering eigen collecties	x	x
1.2. advies, bemiddeling bij verwerven van externe (secundaire) data, beschikbaar stellen daarvan	x	x
1.3. advies over (meta)data		x
2.1. verbeteren beeldbank	x	x
2.2. platform om data in op te slaan, te analyseren en te bewerken		x
2.3. repository voor tools	x	
3.1. tools en technieken		x
3.2. verbinden van onderzoekers		x
3.3. crowdsourcing	x	x

[Dienstverlening voor groepen uit innovatie life cycle]

5. Prioriteiten en roadmap

5.1 Prioriteiten op basis van enquête

5.1.1 Doel en verspreiding

Om preciezer bij digital humanists de behoefte aan ondersteuning te peilen is een enquête uitgezet. Vanwege de internationale samenstelling van sommige onderzoeksgroepen is de enquête in het Engels aangeboden. De enquête is per mail verspreid aan 46 personen. Deze selectie bestond uit mensen met wie gesproken was over digital humanities en de UB, aangevuld tot complete onderzoeksgroepen waar dat relevant leek (bijv. alle onderzoekers van de groep van Piek Vossen). In de mail werd ook gevraagd om de enquête onder de aandacht te brengen van andere geïnteresseerden. Daarnaast is een nieuwsbericht over de enquête geplaatst op:

- de website van het Centre for Digital Humanities
- de website van de UBVU
- VUnet (voor de faculteiten Geesteswetenschappen, Godgeleerdheid, Sociale Wetenschappen en Exacte Wetenschappen)
- UBVU Twitter

Om de respons vanuit computerwetenschappers te verhogen, is door Victor de Boer aan zijn directe collega's per mail een herinnering verstuurd.

5.1.2 Respondenten

In totaal zijn veertien formulieren ingevuld. Twaalf daarvan bevatten resultaten. Van de twee overige is één UB-intern, de andere is niet te herleiden.

In de enquête zijn ook een aantal vragen gesteld over de respondenten zelf. Op basis daarvan kunnen we het volgende zeggen:

<i>Faculteit</i>	<i>Aantal</i>
Humanities	6
Theology	3
Sciences	2
Social Sciences	1
<i>Discipline</i>	<i>Aantal</i>
history	5
linguistics	3
computer science	2
philosophy	1
<i>Expertise DH</i>	<i>Aantal</i>
novice	2
intermediate	6
expert	4

5.1.3 Resultaten

De eigenlijke enquête bestond uit twee delen:

1. vormen van potentiële dienstverlening
2. gebruik van de Beeldbank en van webexposities

5.1.3.1 Potentiële vormen van dienstverlening

Bij iedere vorm van potentiële dienstverlening is gevraagd om op een vijfpuntsschaal aan te geven in hoeverre men het eens was met de stellingen 'I personally have a need for this.', 'I think others might have a need for this.' en 'This is a service that the library should provide.' Elk antwoord is vervolgens omgezet in een aantal punten volgens deze berekening:

Oordeel	Punten
Completely disagree	-2
Disagree	-1
Neutral	0
Agree	+1
Completely agree	+2

Dat levert het volgende op:

	Punten (zelf)	Rangorde (zelf)	Punten (andere)	Rangorde (anderen)
An overview of Digital Humanities projects, tools used at VU University	15	1	18	1
An infrastructure to ensure tools remain available (e.g. in a repository)	14	2	17	3
Advice about or help in acquiring licenses to access data	12	3	17	3
Training in tools or techniques (e.g. Python)	12	3	14	6
A platform to store, analyse and process data	11	5	17	3
Digitisation of material from University Library collections	11	5	18	1
Make secondary data available	9	7	11	7
Advice on (meta)data	7	8	11	7
Opportunities to present projects or project results	6	9	7	9
Contribute to crowdsourcing by mobilising volunteers	5	10	4	10

De laatste vier onderwerpen zijn in beide lijsten het zelfde. In de eerste zes is er wel een duidelijke verschil: behoefte aan digitalisering is bij respondenten zelf gemiddeld aanwezig, maar de inschatting is wel dat het voor anderen heel belangrijk is. Voor de andere onderwerpen in de top-6 zijn de verschillen minimaal.

Men is ook gevraagd aan te geven in hoeverre men het een taak van de UB vindt om dergelijke dienstverlening aan te bieden. Op dezelfde manier gescoord leidt dat tot het volgende overzicht:

	<i>Punten</i>
Digitisation of material from University Library collections	19
Advice about or help in acquiring licenses to access data	14
Advice on (meta)data	14
Make secondary data available	11
An overview of Digital Humanities projects, tools used at VU University	11
A platform to store, analyse and process data	10
An infrastructure to ensure tools remain available (e.g. in a repository)	8
Training in tools or techniques (e.g. Python)	7
Opportunities to present projects or project results	4
Contribute to crowdsourcing by mobilising volunteers	-1

Als we de scores voor eigen behoefte afzetten tegen die voor taak van de UB, komen we tot de volgende visualisatie:



Hierin zijn vijf groepen te herkennen:

1. "Opportunities to present projects or project results" en "Contribute to crowdsourcing by mobilising volunteers"
2. "Make secondary data available" en "A platform to store, analyse and process data"
3. "Training in tools or techniques (e.g. Python)", "An infrastructure to ensure tools remain available (e.g. in a repository)" en "An overview of Digital Humanities projects, tools used at VU University"
4. "Advice on (meta)data"
5. "Digitisation of material from University Library collections" en "Advice about or help in acquiring licenses to access data"

Voor deze vijf categorieën stel ik de volgende prioritering voor:

1. niet actief vervolgen
2. activiteiten van anderen volgen, waar kansen zijn onze expertise inbrengen
3. actief geconstateerde behoefte onder de aandacht brengen bij andere stakeholders en in samenwerking uitwerken op welke manier aan de behoefte voldaan kan worden
4. expertise verder ontwikkelen
5. initiatieven nemen om projecten op te starten

5.1.3.2 Beeldbank en webexposities

Van de respondenten hadden drie gebruik gemaakt van de Beeldbank. Twee van hen gaven ook suggesties ter verbetering:

- "Various ways, too complicated to summarize here." (helaas heeft deze respondent geen mailadres ingevuld, zodat het nog enig speurwerk zal vergen om deze ideeën te achterhalen)
- "It might benefit from a more image first approach presentation (browse through the images)."

Ook een van degenen die aangaf geen gebruik te hebben gemaakt van de beeldbank leverde een suggestie voor verbetering:

- "make it available via europeana.eu"

Metadata van de beeldbankcollecties worden op dit moment geharvest in europeana.eu. In het project Europeana Cloud wordt gewerkt aan onder andere een cloud-infrastructure voor de content zelf en nieuwe diensten en tools voor onderzoekers.¹⁵

De helft van de respondenten vond het ten minste in enige waarschijnlijk dat zij het maken van een webexpositie in hun onderwijs zouden gebruiken.

¹⁵<http://pro.europeana.eu/europeana-cloud/about-europeana-cloud>.

5.2 Roadmap

<i>onderwerp</i>	<i>aanpak</i>	<i>korte termijn</i>		<i>fase 1</i>		<i>fase 2</i>	<i>fase 3</i>
		<i>actie</i>	<i>prioritering</i>	<i>actie</i>	<i>prioritering</i>	<i>actie</i>	<i>actie</i>
4.2.1.1 <i>digitalisering op verzoek</i>	initiatieven nemen om projecten op te starten	samen met conservatoren inventariseren welk materiaal interessant kan zijn voor deze vormen van onderzoek	a	opstellen productplan voor digitalisering op verzoek (waarin beschreven wordt: indicatie van kosten, werkwijze) (afhankelijk van vraag) meewerken aan onderzoeksproject van Fred van Lieburg (werktitel 'VITAE', over religieuze (auto)biografieën van protestanten, 16e-21e eeuw)	a		
4.2.1.2 <i>advies, bemiddeling bij verwerven van externe</i>	initiatieven nemen om projecten op te starten; activiteiten van anderen volgen,	expertmeeting organiseren met onderzoekers die op dit moment vragen hebben gesteld over ondersteuning bij	a	pilot met licentie-onderhandeling (bijv. bij groep van Piek Vossen) belang van TDM	a	pilot met beschikbaar stellen secundaire data van één onderzoeksgroep (bijv. van Piek Vossen)	implementeren van VU-brede oplossing voor secundaire data

<i>(secundaire) data; beschikbaar stellen daarvan</i>	waar kansen zijn onze expertise inbrengen	TDM		integreren in voorlichting over Open Access	a	pilot met adviseren, bemiddelen bij extra rekencapaciteit	
4.2.1.3 advies over (meta)data	expertise verder ontwikkelen	pilot met metadata uit INVENi2	c	aanbieden BNPP als Linked Open Data	b		
4.2.2.1 verbeteren Beeldbank	initiatieven nemen om projecten op te starten	samenstellen van overzicht collectie in Beeldbank, incl. auteursrechtelijke status; informatie hierover toevoegen	b	onderzoek naar mogelijkheden van Europeana, inclusief verbetering van informatie daarover binnen bijv. de Beeldbank en op de website	b		
4.2.2.2 platform om data in op te slaan, te analyseren en te bewerken	activiteiten van anderen volgen, waar kansen zijn onze expertise inbrengen			adviseren bij opstart project voor eLab NLP	c		
4.2.2.3 repository voor tools (of andere resultaten van onderzoek	actief geconstateerde behoefte onder de aandacht brengen bij andere stakeholders en			onderzoek naar behoefte, oplossingsrichtingen en samenwerking met anderen (zoals DANS) in samenwerking met	b		

<i>naast data en publicaties)</i>	in samenwerking uitwerken op welke manier aan de behoefte voldaan kan worden			Centre for Digital Humanities			
<i>4.2.3.1 tools en technieken</i>	actief geconstateerde behoefte onder de aandacht brengen bij andere stakeholders en in samenwerking uitwerken op welke manier aan de behoefte voldaan kan worden	presentatie over ondersteuning bij crash course presentatie over digitalisering bij minor maken van Libguide met o.a. overzicht van veelgebruikte tools	a a b	opzetten van dienst voor literatuuronderzoek voor digital humanities in samenwerking met bijv. Centre for Digital Humanities organiseren van trainingen	b c	ontwikkelen van onderwijsmodule 'tool criticism' DH verwerken in aanbod cursussen informatievaardigheden	
<i>4.2.3.2 verbinden van onderzoekers</i>	actief geconstateerde behoefte onder de aandacht brengen bij andere stakeholders en in samenwerking uitwerken op welke manier	pilot van expertmeeting (digitale) edities	b	in samenwerking met bijv. Centre for Digital Humanities opstellen van overzicht van projecten bij UvA/VU	c	op basis van informatie in CRIS en repository bijdragen aan expertise-overzicht bij de VU, o.a. op het gebied van digital humanities	

	aan de behoefte voldaan kan worden						
<i>organisatie- ontwikkeling</i>		samenstelling DH- team vak/themaspecialisten met kennis van visie	a a	vak/themaspecialisten met voldoende kennis voor herkennen kansen in onderwijs/onderzoek	b		

Bijlagen

Concept 'digital humanities'

'e-humanities' of 'digital humanities'

De termen 'e-humanities' en 'digital humanities' worden vaak inwisselbaar gebruikt. Ook binnen de organisatie van de VU komen beide termen voor: enerzijds heeft de minor van de FGW als titel 'Digital Humanities'¹⁶, anderzijds heet het keuzepakket in BA2 van de opleiding Geschiedenis 'E-humanities'¹⁷. Serge ter Braake (FGW) schreef er ook een blogpost¹⁸ over, met als conclusie dat semantisch de term 'e-humanities' misschien rijker is, maar dat de term 'digital humanities' te verkiezen is, omdat die term meer gebruikt wordt. In deze verkenning gebruiken we daarom ook de term 'digital humanities'.

In het vervolg worden degenen die zich bezig houden met digital humanities, zowel onderzoekers als studenten, aangeduid met de koepelterm 'digital humanists'.

Definitie van digital humanities

Nicoline van der Sijs geeft de volgende definitie van digital humanities¹⁹: “een (...) vakgebied dat tracht met behulp van computertechnieken antwoorden te vinden op geesteswetenschappelijke onderzoeksvragen”.

Soms wordt onderscheid gemaakt tussen twee verschijningsvormen van digital humanities, zoals door Stephen Ramsay²⁰:

1. “set of practices that most regarded as intimately related: text encoding, archive creation, text analysis, historical gis, 3d modeling of archaeological sites, art historical cataloging, visualization, and general meditation on what all of these new affordances might mean for the study of the human record”
2. “humanistic inquiry that in some way relates to the digital”: “Media studies practitioners were digital humanists; people who had devoted several decades to digital pedagogy were digital humanists; cultural critics who were interested in Internet culture were digital humanists; and digital artists of a certain variety were digital humanists.”

Het Office of Digital Humanities van de National Endowment for the Humanities maakt hetzelfde onderscheid²¹:

1. “projects that explore how to harness new technology for humanities research”
2. “those that study digital culture from a humanistic perspective”

Een kenmerkend verschil tussen de eerste en tweede verschijningsvorm is dat bij de tweede het onderzoeksmateriaal uit zichzelf al digitaal is, terwijl voor de eerste verschijningsvorm analoog materiaal gedigitaliseerd wordt. De tweede verschijningsvorm komt bij de VU bijvoorbeeld voor bij Kunstgeschiedenis (post-digital art) of Communicatiewetenschap. De eerste verschijningsvorm komt vaker voor, bijv. bij Geschiedenis of Godgeleerdheid. In deze verkenning wordt rekening gehouden met beide verschijningsvormen.

¹⁶<http://fgw.vu.nl/nl/opleidingen/minoren/minor-digital-humanities/index.asp>

¹⁷<http://www.vu.nl/nl/studiegids/2014-2015/bachelor/g-k/geschiedenis/index.asp?view=modulegroep&origin=50975769&id=50980211>

¹⁸<http://sergeterbraake.blogspot.nl/2014/02/e-humanities-or-digital-humanities.html>

¹⁹<https://www.ensie.nl/nicoline-van-der-sijs/e-humanities>

²⁰<http://stephenramsay.us/2013/05/03/dh-one-and-two/>

²¹<http://www.neh.gov/divisions/odh/about>

Een goed overzicht van activiteiten die horen bij digital humanities is de ‘taxonomy of digital research activities in the humanities’ (TaDiRAH)²².

Digital Humanities bij andere bibliotheken

dh+lib BeNeLux

Op 30 maart vond in de KB een bijeenkomst plaats voor diegenen die zich bij een academische bibliotheek in Nederland of Vlaanderen bezighouden met digital humanities of die in dat onderwerp geïnteresseerd waren. De bijeenkomst was een initiatief van de KB (Steven Claeysens), de Letterenbibliotheek Gent (Saskia Scheltjens) en de VU (Michiel Cock). Bij die bijeenkomst kwam duidelijk naar voren dat veel bibliotheken zich van de ontwikkelingen bewust zijn, maar dat bij iedereen de dienstverlening nog in ontwikkeling is.

De groep heeft als titel dh+lib gekregen, vernoemd naar het internationale initiatief. In aanloop naar het congres DHBenelux 2015 is besloten om de groep uit te breiden tot ook Luxemburg, en aansluiting te zoeken bij de organisatie van de DHBenelux-congressen. Er wordt in ieder geval een discussielijst ingericht, om kennisdeling tussen de verschillende bibliotheken te bevorderen.

Hieronder volgt een kort overzicht van relevante ontwikkelingen en diensten bij andere bibliotheken.

KB

De KB biedt op verschillende manieren ondersteuning voor digital humanists²³:

1. data, zoals bijvoorbeeld Delpher
2. een KB Research Lab, met daarin tools voor onderzoek met de KB-data
3. programma's waarbij onderzoekers uitgenodigd worden om in de KB en met (technische) ondersteuning onderzoek te doen
4. organisatie van congressen en workshops, zoals DHBenelux 2014, het symposium *Historische kranten als 'big data'*, of een workshop topic modeling met behulp van de tool MALLET

De verschillende vormen van ondersteuning worden gecommuniceerd op een website en door middel van een flyer. Veel nieuws wordt daarnaast ook verspreid via het (algemenere) Twitteraccount @KBNLresearch.

Faculteitsbibliotheek Letteren en Wijsbegeerte Gent: Library Lab

Dit Library Lab²⁴ biedt “tutorials, workshops, lezingen en adviesmomenten” om digital humanities te ondersteunen. Men organiseert bijvoorbeeld een vierdaagse cursus Python en gelegenheden waarbij projecten zichzelf kunnen presenteren. De onderwerpen van dit Library Lab zijn ruim: ook auteursrecht, Twitter en Zotero komen aan bod. Ook de lezing die Wouter Gerritsma in maart in Gent gaf over Social Media en Altmetrics vond plaats binnen het kader van het Library Lab.

Universitaire Bibliotheken Leiden

Bij de UBL wordt een Centre for Digital Scholarship opgericht, maar de precieze invulling van dit centrum is op dit nog niet bepaald. Er zijn wel al twee pilotprojecten gedaan (waarvan een bij Science en een bij Geesteswetenschappen), waarin de mogelijkheden van text mining onderzocht werden. Gezien de naam zal het centrum niet alleen digital humanities ondersteunen, maar digitale vernieuwing in alle wetenschappen faciliteren.

²²<http://tadirah.dariah.eu> en <https://github.com/dhtaxonomy/TaDiRAH>

²³<http://www.kb.nl/digitalhumanities>

²⁴<http://www.flw.ugent.be/bibliotheek/librarylab>

Internationale ontwikkelingen (buiten de Benelux)

De dh+lib Registry²⁵ biedt een overzicht van ondersteuning bij verschillende bibliotheken in met name de VS.

²⁵<http://acrl.ala.org/dh/registry/>