SUPPORTING EXPLORATORY SEARCH WITH FEATURES, VISUALIZATIONS, AND INTERFACE DESIGN: A THEORETICAL FRAMEWORK.

Maartje Kruijt

maartjekruijt@hotmail.com
10430563

17 juni 2016
Begeleiders Jan Simons en Victor de Boer
Media en Cultuur
Universiteit van Amsterdam

ABSTRACT

In de hedendaagse netwerksamenleving is een groeiende behoefte aan het delen, integreren en zoeken binnen collecties van verschillende bibliotheken, archieven en musea (Brown 4; Evens en Hauttekeete 163). Wetenschappers hebben *tools* nodig die interpretatie van deze collecties ondersteunen. In de exploratieve fase van onderzoek heeft de onderzoeker geen duidelijke focus en is afwisselend onzeker, het is van belang om zoeksystemen te ontwikkelen die de onderzoeker hierbij ondersteunen. Datavisualisaties worden geprezen om het vermogen om verhalen in data bloot te leggen. Deze cognitieve eigenschappen lijken strategieën en tactieken te ondersteunen die van belang zijn in het doen van exploratief onderzoek (Crédeville en Forest 174). Het visualiseren van informatie bij aan het analyseren van patronen in data (Crédeville en Forest 174). Collectieobjecten moeten in een exploratieve fase zichtbaar zijn, visualisaties kunnen bijdragen aan exploratie maar moeten de initiële objecten niet vervangen (Granitzer et al. 127; Hornbæk en Frøkjær 180). Theoretisch gezien kan het implementeren van narratieven in de *linked data* browser DIVE, bijdragen aan het interpretatieproces van onderzoekers. Dit kan zowel autogegenereerd, op basis van bestaande *event-event* relaties, als handmatig door onderzoekers worden gedaan. Onderzoekers de mogelijkheid geven objecten uit de DIVE database in een narratieve vorm te presenteren brengt een aantal voordelen met zich mee. Onderzoekers kunnen zo hun bronnen en onderzoek interactief presenteren. Ook kunnen ze het narratief gebruiken als referentie en delen met andere onderzoekers. Het interactief presenteren van onderzoeksresultaten gaat ten koste van de narrativiteit van een tekst, wel gaat het geheel er in termen van communicatie op vooruit (Allenye 160). De geconstrueerde narratieven kunnen vervolgens dienen als beginpunt voor verder exploratie van andere onderzoekers die gebruik maken van de DIVE browser.

DANKWOORD

Bij dezen zou ik graag Victor de Boer willen bedanken voor zijn tijd, geduld en optimisme. Ook zou ik graag Jan Simons bedanken voor tijd en constructieve kritiek. Daarnaast wil ik graag het hele DIVE team bedanken voor de manier waarop ze me hebben verwelkomt in het team en voor de tijd die ze voor mij hebben vrijgemaakt. In het speciaal wil ik graag Liliana Melgar bedanken voor haar steun gedurende het hele proces.

INHOUDSOPGAVE

1. INLEIDING **5**
1.1 ONDERZOEKSVRAAG *8*
1.2 DOEL/DOELSTELLINGEN EN OMVANG *9*
1.3 METHODE *9*
1.4 DIVE *9*
1.5 OPZET VAN SCRIPTIE *10*

2. IMPLICATIES VAN VERSCHILLENDE FASES VAN ONDERZOEK VOOR ZOEKSYSTEMEN **12**
2.1 INFORMATIE ZOEKMODELLEN *13*
2.2 EXPLORATIEF ZOEKEN EN ONDERZOEK *15*
2.3 INTERFACE DESIGN *17*
2.3 IMPACT VAN INFORMATIE ZOEKMODELLEN OP INFORMATIEZOEKSYSTEMEN *19*

3. TECHNOLOGIEËN DIE EXPLORATIEF ONDERZOEK ONDERSTEUNEN **21**
3.1 HET SEMANTISCHE WEB EN LINKED DATA *21*
3.2 IMPLICATIES LINKED DATA VOOR EXPLORATIEVE ZOEKBROWSERS *22*
3.3 EVENTS EN NARRATIEVEN *23*
3.4 TECHNISCHE IMPLEMENTATIE VAN EVENTS EN NARRATIEVEN *24*

4. VISUALISATIES IN EEN EXPLORATIEVE ONDERZOEKSFASE **27**
4.1 NARRATIEVE PRESENTATIES *28*
4.2 DIVE INTERFACE *29*
4.3 IMPLEMENTATIE EN PRESENTATIE VAN NARRATIEVEN IN DIVE *30*

5. CONCLUSIE EN VERVOLGONDERZOEK **34**

6. REFERENTIES **37**

7. APPENDIX 1 **47**
USER STORY 07 "CONSTRUCT & STORE NARRATIVES WHILE NAVIGATING THE DATA" *47*
USER STORY 07A: AS A RESEARCHER I WANT TO SAVE MY SEARCH PATH / NARRATIVE SO I CAN LOOK BACK AT THIS FOR FUTURE REFERENCE 47
USER STORY 07B: AS A RESEARCHER I WANT TO SAVE SPECIFIC OBJECTS I FOUND DURING BROWSING SO I CAN BUILD A CORPUS ON MY CHOSEN TOPIC 47
USER STORY 07C: AS A RESEARCHER I WANT TO SHARE MY SEARCH PATH / NARRATIVE WITH OTHER RESEARCHERS SO MY SEARCH PATH CAN BE USED BY OTHERS 48
USER STORY 09 "LINK TO MY NARRATIVE SO I CAN REFER TO IT IN OTHER SYSTEMS" *48*
USER STORY 10 "PUBLISH MY NARRATIVES & SHARE IT WITH OTHERS" *49*
USER STORY 08 "PLOT MY NARRATIVE(S) ON A TIMELINE" *49*
USER STORY 20 "CREATE COLLECTIONS OF CERTAIN EVENTS AND MEDIA OBJECTS" *50*
USER STORY 21 "UNDERSTAND THE SEMANTICS OF PATHS TO IDENTIFY INTERESTING NARRATIVES" *50*
USER STORY 30: AS A RESEARCHER I WANT TO BE ABLE TO SEE OTHER PEOPLE'S NARRATIVE THAT ARE MOST RELEVANT TO MY SEARCH SO I CAN QUICKLY COME TO INTERESTING AND SURPRISING RESULTS *50*

8. APPENDIX 2 **51**
OVERZICHT VERSCHILLENDE ONDERZOEKSFASES *51*

1. INLEIDING

De beschikbaarheid van digitale informatie hervormt de productie, verspreiding, het gebruik en hergebruik van informatie en kennis in de onze samenleving fundamenteel (Evens en Hauttekeete 158; Ross en Hedstrom 317). Dit heeft ook betrekking op bibliotheken, archieven en musea (LAMs: Libraries Archives and Museums). In de hedendaagse netwerksamenleving is een groeiende behoefte aan het delen, integreren en zoeken binnen collecties van verschillende LAMs (Brown 4; Evens en Hauttekeete 163). Collecties en de bijbehorende data van erfgoedinstellingen moeten digitaal beschikbaar worden gemaakt voor zowel het grote publiek als voor onderzoekers. Het is van belang dat dit gebeurt op een manier dat collecties eenvoudig geintegreerd kunnen worden, bijvoorbeeld door middel van *linked data*.

De manier waarop deze erfgoedinstanties toegang bieden tot hun collecties in het digitale tijdperk moet worden herzien (Van Passel en Rigole 203). Wetenschappers hebben *tools* nodig die interpretatie van collecties ondersteunen. Het is cruciaal dat de culturele erfgoedinstanties hun best doen om er voor te zorgen dat gebruikers niet verdwalen in hun online archief, maar hen de weg wijzen in hun zoektocht naar informatie en interpretatie (De Boer et al. 1; Ross en Hedstrom 317). Veel van de instituten hebben het echter moeilijk met het ondersteunen van hun online bezoekers in het toegang geven tot hun collecties. Bezoekers hebben meer toegang maar begrijpen minder. Dit is voornamelijk te verklaren door ontbrekende curatie (De Boer et al. 1). Gebrek aan context doet de bezoekers verdwalen in de enorme online portals. Dit geldt zowel voor professionele als voor reguliere bezoekers (Van den Akker et al. 'From Information Delivery to Interpretation Support' 431; De Boer et al. 1; Van Passel en Rigole 2).

Ook moeten de instituten nieuwe manieren bedenken waarop gebruikers, zowel professioneel als regulier, hun collecties kunnen verrijken. (Van den Akker et al. 'Digital Hermeneutics' 1; Van den Akker et al. 'From Information Delivery to Interpretation Support' 431; Ross en Hedstrom 317). Het web kan op deze manier worden gezien als een plek waar instituten en hun gebruikers met elkaar in dialoog kunnen treden. Zo'n dialoog kan leiden tot nieuwe inzichten met betrekking tot culturele erfgoedinstellingen, zowel voor het reguliere publiek als wetenschappers (Van den Akker et al. 'Digital Hermeneutics' 7; Van den Akker et al. 'History as Dialogue: On online Narrativity' 113).

Zowel geesteswetenschappers als sociale wetenschappers bestuderen de menselijke cultuur. Volgens José van Dijck zijn geesteswetenschappers uit verschillende disciplines voornamelijk goed in het interpreteren van inhoud van afzonderlijke data op hun eigen terrein (1). Doordat de data waar alfa-wetenschappers mee werken (zoals tekst beeld, geluid en historische data) talrijk en versnipperd zijn, focussen de meeste geesteswetenschappers zich op een klein puzzelstukje. Waar wetenschappers uit deze disciplines tot nu toe nog minder goed in zijn geslaagd is de verschillende vragen in een omvattend verband te plaatsen (van Dijck 1).

De beschikbaarheid van digitale data van erfgoedinstellingen brengen zowel nieuwe mogelijkheden als uitdagingen met zich mee. Allereerst zijn er instrumenten nodig voor onderzoekers om deze enorme hoeveelheden aan data te bevragen en interpreteren. Dit verandert zowel de manier waarop geesteswetenschappers onderzoek doen als de vragen die zij stellen en hoe ze dezen beantwoorden (van Dijck 1; Borgman 8; Collins en Micheal 138; Michel et al. 1; Toms en O'Brien 102). Bron et al. geven aan dat als we *tools* willen ontwikkelen die onderzoek in de geesteswetenschappen (en sociale wetenschappen) pogen te ondersteunen, er een gedetailleerd begrip nodig is van de relatie tussen de verschillende databronnen, *tools* en de activiteiten in het onderzoeksproces. Het is bovendien van belang om te kijken naar hoe deze interacties invloed hebben op de onderzoeksvraag. (1535).

In dit onderzoek wordt onderzocht hoe wetenschappers, in het bijzonder historici, in de exploratieve fase van onderzoek kunnen worden ondersteund bij het contextualiseren en interpreteren van historische objecten door middel van visualisaties/features in interactieve zoekbrowsers. Dit onderzoek vindt plaats in de context van het *linked data* project DIVE. DIVE is een *event* gebasseerde *linked data* zoekbrowser die openbare data uit musea en bibliotheken toegankelijk en navigeerbaar maakt voor onderzoekers, studenten en het grote publiek. Het DIVE project is onderdeel van *Common Lab Research Infrastructure for the Arts and Humanities*[1] (CLARIAH). DIVE is ontwikkeld om interactieve toegang te bieden tot historische objecten uit heterogene

[1] CLARIAH houdt zich bezig met het maken van duurzame en gebruiksvriendelijke digitale applicaties voor onderzoekers in de geestes- en sociale wetenschappen om zo vernieuwend geesteswetenschappelijk onderzoek mogelijk te maken (CLARIAH website).

collecties door middel van historische *events* en narratieven te gebruiken als context voor zoeken, browsen en het groeperen van objecten (De Boer et al. 1-2).

Events bevatten informatie over een persoon, plaats, periode of een type die geassocieerd worden met een bepaald object. *Events* kunnen naast *object-event* relaties ook *event-event* relaties hebben. Het project maakt gebruik van *digital hermeneutics* om relaties tussen historischen *events* en objecten te laten zien. De voornaamste doelstelling van *digital hermeneutics* is om de relatie tussen het menselijke interpretatie proces en web applicaties die dit proces ondersteunen te onderzoeken (Van den Akker et al. 'Digital Hermeneutics' 1; Capurro 35). Door historische objecten in een *linked data* netwerk te visualiseren poogt DIVE onderzoekers te ondersteunen in de exploratieve fase van onderzoek en het formuleren van een onderzoeksvraag. DIVE kan hierdoor worden gezien als een heuristische onderzoekstool.

Onderzoek van Van Erp et al. geeft aan dat veel gebruikers behoefte hebben aan meer exploratieve vormen van browsen binnen online erfgoedcollecties (1). Voorgaand onderzoek met het Agora project, de voorloper van de DIVE tool, laat zien dat interactievere manieren van browsen door deze collecties als behulpzaam kan worden ervaren voor het interpretatieproces van de gebruiker (Van den Akker et al. 'Digital Hermeneutics' 4). Ook bleek het zichtbaar maken van narratieve relaties bij te dragen aan het contextualiseren en interpreteren van data uit cultureel erfgoedcollecties (Van den Akker et al. 'Digital Hermeneutics' 7). Dit bleek onder meer uit verschillende user studies waarin onderzoekers aangaven behoefte te hebben om "[to] Construct & store narratives while navigating the data" (*Appendix 1*: *user story* 7). Dit was om verschillende redenen bijvoorbeeld voor "future reference"(*Appendix 1*: *user story* 7a) of "[to] Link to my narrative so I can refer to it in other systems" (*Appendix 1*: *user story* 9). Daarbovenop wilden onderzoekers graag de mogelijkheid hebben om "[to] publish my narratives and share it with others"(*Appendix 1*: *user story* 7c; *user story* 10). Als laatste gaven onderzoekers aan dat ze graag andermans narratieven wilden zien die relevant waren aan hun zoekopdracht "So I can quickly come to interesting and surprising results" (*Appendix 1*: *user story* 30).

Aangezien *events* centraal staan in DIVE, en *events* een van de bouwstenen van een narratief zijn, kan dit een goed beginpunt bieden voor het implementeren van narratieven in DIVE (*Appendix 1*: *user story* 7). Het vertellen en creëren van verhalen maar ook de visuele expressie hiervan, zijn integrale onderdelen van de menselijke

cultuur (Segel en Heer 1; Alleyne 1). Narratieven mediëren het verleden en de wereld om ons heen. Ons begrip van de wereld wordt door middel van *events* en gebeurtenissen overgedragen aan anderen, hierin zijn objecten en abstracte noties gegrond in ruimte en tijd door hun deelname aan gebeurtenissen (Van den Akker et al. 'From Information Delivery to Interpretation Support' 432).

Historici vertellen verhalen over gebeurtenissen uit het verleden. Deze gebeurtenissen plaatsen ze in een bepaalde narratieve context waarmee ze tot een eigen presentatie en interpretatie van deze gebeurtenissen komen. Historici analyseren op deze manier overblijfselen van het verleden om te achterhalen wat er is gebeurd, waarom gebeurtenissen hebben plaatsgevonden en hoe het heeft bijgedragen aan de wereld zoals die vandaag de dag is. Het gebruik van materiele objecten in historisch onderzoek is niet evident. Traditioneel gezien maken historici gebruik van geschreven bronnen. Cultuurhistorici, archeologen en kunstgeschiedenis wetenschapper zijn hier een uitzondering op en bestuderen wel materiele objecten. Aangezien DIVE een *linked data* browser is met historische objecten (en hierbij behorende *events*, concepten, plaatsen en personen) zal dit onderzoek zich richten op historici die wel materiele objecten bestuderen.

Dit onderzoek richt zich op hoe zoekbrowsers onderzoekers kunnen ondersteunen in een exploratieve fase van onderzoek. Specifiek wordt ingegaan op de vraag of presentaties van narratieven, onderzoekers kunnen ondersteunen bij het contextualiseren en interpreteren van historische objecten die beschikbaar zijn in de *linked data tool* DIVE. Maar context kan op verschillende manieren worden construereerd en gepresenteerd en interpretatie kan eveneens op meerdere manieren ondersteund worden. Daarom worden ook andere technologieen besproken die de exploratieve onderzoeksfase kunnen ondersteunen.

1.1 Onderzoeksvraag

Hoe kunnen wetenschappers, in het bijzonder historici, in de exploratieve fase van onderzoek kunnen worden ondersteund bij het contextualiseren en interpreteren van historische collecties door middel van features in interactieve zoekbrowsers, en hoe kunnen narratieve presentaties hieraan bijdragen?

1.2 DOELSTELLINGEN EN OMVANG

Mijn bijdrage aan dit project bestaat uit het ontwikkelen van een beknopt overzicht van theorieën en technologieën die een exploratieve fase van onderzoek ondersteunen. Ook worden aanbevelingen gedaan voor verbeteringen in de interface die ten dienste staan van onderzoekers in het interpreteren en contextualiseren van online ergoedcollecties.

1.3 METHODE

Het voornaamste deel van dit onderzoek bestaat uit kwalitatief literatuuronderzoek. Dit heeft als doel een breder begrip van belangrijke concepten te creëren en het situeren van het onderzoek in bestaande literatuur. Het literatuurverslag biedt een kader van het onderzochte probleem (Creswell 6). Het literatuurreview kan daarnaast ook leiden tot methodologische inzichten (Gall, Borg en Gall 2). We valideren de bevindingen van het onderzoek met het ontwikkelingsteam om tot een nieuw prototype te komen.

1.4 DIVE

DIVE is een *linked data* browser die het mogelijk maakt objecten uit verschillende heterogene culturele erfgoedcollecties aan elkaar te verbinden door middel van historische *events* en narratieven (De Boer et al. 1-2). In de DIVE *triple store*[2] zijn op het moment vier collecties te vinden. Deze bestaan uit 3000 video's van Openbeelden[3] waarvan 300 video's verrijkt zijn door middel van *crowdsourcing*. Daarnaast enkele miljoenen Radio Bulletins te vinden afkomstig uit de Koninklijke bibliotheek[4] - hiervan is een deelverzameling van 2210 gedigitaliseerde typoscripten te vinden. Deze bestaat onder andere uit radio en nieuws-scripts uit de periode 1937 tot 1984. Deze verzameling in geselecteerd op basis van de OpenBeelden selectie. Daarnaast zijn er duizenden objecten van het Amsterdam Museum[5] te vinden. Hier is ook een *sub set* van aangemaakt die overeenkomt met de dataset van OpenBeelden. Items uit deze deel set

[2] data.dive.beeldengeluid.nl

[3] www.openbeelden.nl

[4] Deze zijn origineel afkomstig van delpher.kb.nl

[5] Origineel afkomstig van https://bitbucket.org/biktorrr/amlod

dateren van 1950 tot 1980. In het totaal zijn er nu 3576 objecten met 140K *triples* . Van het Tropenmuseum is er een dataset van 70.000 objecten[6]. Uit deze dataset is een selectie gemaakt gebaseerd op een 'serie' (wat te vergelijken is met een collectie). Deze serie is door het Tropenmuseum zelf samengesteld. De willekeurig geselecteerde serie heeft ongeveer 1000 objecten, die allemaal bestaan uit foto's.

De *tool* maakt gebruik van de semantiek uit de bestaande collectievocabulaire en *linked data*-vocabulaire om connecties te maken tussen objecten in de collectie en *events*, personen, plaatsen, en concepten die geassocieerd worden met deze objecten (De Boer et al. 1-2). Dit netwerk van *events*, personen, plaatsen en concepten biedt context aan de culturele erfgoedobjecten die worden gerepresenteerd in hetzelfde netwerk, dit is fundamenteel voor de vindbaarheid en hermeneutiek van de objecten die zich hierin bevinden (De Boer et al. 2). Op deze manier ondersteund DIVE *digital humanities* onderzoekers en studenten in hun online verkenning en in het formuleren van een onderzoeksvraag door middel van contextualisatie van historische objecten. (De Boer et al. 1-2). Dit is in tegenstelling tot andere *digital humanities tools* die zich voornamelijk richten op het beantwoorden van een onderzoeksvraag.

1.5 Opzet van scriptie

Het eerste deel van dit onderzoek gaat in op de verschillende fases van onderzoek. Hierin wordt expliciet gekeken naar de exploratieve fase van onderzoek, de activitieten en informatiebehoeften van onderzoekers in deze fase en vervolgens naar de eventuele implicaties die deze hebben voor zoekbrowsers en interfaces. Ook wordt in het tweede hoofdstuk gekeken naar relevante literatuur op het gebied van interface design.

Vervolgens wordt gekeken naar technologieen die het interpretatieproces van de gebruiker kunnen ondersteunen, zoals semantische annotatie en linked data. Ook wordt de rol van events en narratieven hierbinnen besproken.

Het vierde hoofdstuk van dit onderzoek wordt gekeken hoe informatievisualisaties bij kunnen dragen aan het interpreteren van een erfgoedcollectie in een exploratieve fase van onderzoek. Ook wordt het standpunt van Van den Akker onderzocht en wordt gekeken naar hoe narratieve relaties tussen *events*, objecten,

[6] Deze zijn afkomstig van eculture.multimedian.nl

personen, plaatsen en concepten in online beschikbare cultureel erfgoedcollecties context kunnen bieden in exploratieve fase van onderzoek. De huidige interface van DIVE wordt besproken en aan de hand van uitkomsten uit de voorgaande hoofdstukken zal een voorstel worden gedaan voor verbetering van de huidige DIVE interface.

2. IMPLICATIES VAN VERSCHILLENDE FASES VAN ONDERZOEK VOOR ZOEKSYSTEMEN

Het doen van onderzoek is in verschillende fases te onderscheiden (Marchionini 'Exploratory Search' 41; Kuhlthau 'Kuhlthau's information search process's 369; Vakkari 'A theory of the task-based information retrieval process' 48; Huurdeman en Kamps 3). In de verschillende fases van onderzoek hebben wetenschappers verschillende onderzoeksactiviteiten en informatiebehoeftes; dit heeft implicaties voor de interactie met deze zoeksystemen (Huurdeman en Kamps 3; Vakkari 'Task complexity, problem structure and information actions' 819; Belkin, Seeger en Wersig 351; Checkland en Holwell 5; Ingwersen 3; Järvelin 35). Daarbij kan informatievergaring op talloze manieren worden beschreven afhankelijk van het perspectief en is daarom geformuleerd in een grote hoeveelheid modellen (Wilson 252). Voorbeelden hiervan zijn Kuhlthau's *Information Seeking Process* (ISP), Vakkari's *Task Based Information Retrieval* (IR) *Process* en het in *Appendix 2* geformuleerde schema door Melgar[7] .

Veel van de huidige zoeksystemen hanteren een "one-size fits all" aanpak: er wordt gebruikt gemaakt van één interface gedurende alle verschillende fases van onderzoek (Huurdeman en Kamps 2). Daardoor kan het zo zijn dat het zoekproces van de gebruiker niet adequaat wordt ondersteund in de context van complexe zoektaken (Brusilovsky en Maybury 31; Hearst 'Search user interfaces' 173; Ingwersen en Järvelin 1; Kamps 2; White en Roth 3; Wilson et al. 5). Huurdeman en Kamps onderzoeken de implicaties van verschillende fases van onderzoek voor zoeksystemen (3). Onderzoek over informatiezoeksystemen is tot nu toe niet heel concreet geweest in het discussiëren van tastbare ondersteuning voor verschillende zoekfases (Huurdeman en Kamps 3).

Daarom wordt gekeken naar Interface design voor zoeksystemen, dit is een belangrijk probleem voor *information retrieval* in het algemeen (Walker 14). Sinds eind

[7] Andere voorbeelden hiervan zijn Elis's *behavioral model* (171), Foster's non-lineaire model van informatie zoeken (255), Dervin's *Sense Making* theorie (9-10) en Wilson's probleem oplossende model (266). Al de voorgenoemde modellen benadrukken andere aspecten van zoeker naar informatie en hebben een eigen aanpak en *framework* voor het denken over problemen (Huurdeman en Kamps 3).

jaren negentig is interface design een actief gebied van onderzoek binnen *information retrieval* geworden (Belkin et al. 404). Het meeste onderzoek in het gebied van informatiezoeksystemen heeft zicht gericht op systeemfunctionaliteiten. Interfacevraagstukken werden genegeerd, wat leidt tot een situatie waarin een interface wordt gezien als iets dat bovenop een *information retrieval* systeem wordt geplakt, in plaats van hier een integraal onderdeel van te zien (Marchionini en Komlodi 89).

In dit hoofdstuk wordt een literatuuroverzicht gegeven met betrekking tot de onderverdeling van onderzoekfases. Hieruit volgt een duidelijke formulering van problemen waar onderzoekers tegenaanlopen in exploratieve fases van onderzoek. Aan de hand van interface designtheorie wordt gezocht naar bestaande oplossingen voor deze problemen. Het uiteindelijke doel is om de kloof te overwinnen tussen macro modellen die fases van informatie zoeken in kaart brengen en hoe deze op micro niveau ondersteund kunnen worden door zoeksystemen en interfaces.

2.1 INFORMATIE ZOEKMODELLEN

Kuhlthau's ISP presenteert het informatiezoekproces als een constructie (361). Ze onderscheidt zes verschillende onderzoeksfases: *initiation, selection, exploration, formulation, collection and presentation* (369). Ze formuleerde haar model rond 1980 en verfijnde het rond 1990. Recenter onderzoek wijst uit dat haar model nog steeds bruikbaar is in hedendaagse technologisch rijke informatieomgevingen als het gaat om het ontwerpen van gebruiker gecentreerde informatiesystemen (Kuhlthau, Heinström en Todd 11). Vernieuwend aan haar model was dat ze rekening hield met affectieve aspecten samen met cognitieve en fysieke aspecten, zoals te zien is in figuur 1 (Huurdeman en Kamps 2). Volgens Kuhlthau, is onzekerheid een van de kernconcepten in het proces van het zoeken naar informatie. Onzekerheid is volgens Kuhlthau onderworpen aan fluctuatie tijdens verschillende onderzoeksfases. Zoals in haar model te zien is, zijn onderzoekers vooral in de eerste drie fases wisselend onzeker of optimistisch, in de latere fases van onderzoek neemt deze onzekerheid af maar blijft wel een rol spelen. Ze stelt dat hierop geanticipeerd dient te worden middels verbeteringen in het aanbod van informatiesystemen in eerdere vormende fases van onderzoek (369-70).

Vakkari bouwt voort op het model van Kuhlthau. Hij heeft de verschillende onderzoeksfases echter onderverdeeld in drie categorieën*: pre-focus* (dit kan worden vergleken met fase een, twee en drie van Kuhlthau), *focus-formulering* (fase vier) en *post-focus* (fase vijf en zes) ('A theory of the task-based information retrieval process' 48). Vakkari benadrukt de cruciale rol van het vinden van een focus in zoekprocessen. In de *pre-focus* fase, zijn gedachtes volgens hem "generaal, gefragmenteerd en vaag" (Vakkari 'A theory of the task-based information retrieval process' 46). In deze fase is het moeilijk voor een onderzoeker om concreet te uiten wat voor informatie zij nodig heeft. Dit heeft bijvoorbeeld als gevolg dat in het begin van het zoekproces, het vermogen van de onderzoeker om in te schatten of een bron belangrijk is of niet, erg slecht is. Hij verklaart dit door de ongestructureerde mentale representatie van het onderwerp (Huurdeman en Kamps 4). Als onderzoekers minder kennis over een bepaald onderwerp hebben zullen bronnen eerder als relevant worden beoordeeld (Pharo en Nordlie 9; Spink, Greisdorf en Bateman 600; Vakkari en Hakala 541; Vakkari, Pennanen en Serola 447). Na het vinden van een focus is het zoeken gerichter, wat leidt tot meer relevante informatie (Vakkari 'A theory of the task-based information retrieval process' 46). Zowel Vakkari als Kuhlthau, demonstreerde dat de *gezochte informatie*, de *relevantie* en de *zoektactiek en termen en operaties* varieerden gedurende verschillende fases in onderzoek (Huurdeman en Kamps 1).

In *Appendix 2* is een model te vinden van verschillende onderzoeksfases geformuleerd door Melgar. Melgar bouwt voort op Bron et al. en concepten van Huurdeman en Kamps (met daarin de bovengenoemde modellen van Kuhlthau en Vakkari) en De Jong et al. Dit model zal in dit onderzoek worden gebruikt als basis voor de verschillende onderzoeksfases. In het model wordt een primair onderscheid gemaakt tussen vier verschillende onderzoeksfases. Allereerst de exploratieve onderzoeksfase die te vergelijken is met fase een, twee en drie van Kuhlthau en de *pre-focus* fase van Kuhlthau. Daarnaast maakt Melgar onderscheid tussen de samenstellende en analyse fase. Volgens Melgar zijn deze fases zo goed als gelijk aan elkaar en vindt er een continue iteratie plaats tussen de twee fases. Deze fases kunnen worden vergeleken met fase vier en vijf uit Kuhlthau's model en Vakkari's *focus-formulering*. De laatste fase die Melgar onderscheidt is de presentatie fase van onderzoek, deze fase komt overeen met fase zes uit Kuhlthau model en de *post-focus* fase van Vakkari.

In het model van Melgar zijn de belangrijkste concepten uit de modellen van Kuhlthau en Vakkari terug te vinden. Bijvoorbeeld onder het kopje "feelings/cognition" is te zien dat onderzoekers verward, onzeker en twijfelen in de exploratieve fase van onderzoek (*Appendix 2*). Kuhlthau geeft aan dat in de *pre-focus* fase van onderzoek de onderzoeker geen duidelijke focus heeft en moeite heeft met het formuleren van haar informatiebehoefte (Kuhlthau 'Seeking meaning' 41). Zowel Melgar als Kuhlthau geven aan dat dit opgelost wordt op het moment dat de onderzoeker een duidelijke probleem formulering heeft: gebruiker "[...] ends up with formulation: clarity" (Melgar *Appendix 2*). Voor een duidelijke probleemformulering heeft de onderzoeker bepaalde ondersteuning nodig gedurende zijn zoekproces. In de volgende paragraaf wordt ingegaan op de methode en informatiebehoefte in de exploratieve onderzoeksfase.

Figuur 1 Kuhlthau's zoekfases

Stage	*Description*
1. Initiation	becoming aware of a lack of knowledge or understanding, often causing uncertainty
2. Selection	identifying & selecting general area, topic or problem, sense of optimism replaces uncertainty
3. Exploration	exploring & seeking information on the general topic, inconsistent info can cause uncertainty
4. Formulation	focused perspective is formed, uncertainty is reducing, while confidence increases
5. Collection	gathering pertinent information to focused topic, less uncertainty, more interest/involvement
6. Presentation	completing the search, reporting and using results

Kuhlthau 'Seeking meaning' 82

2.2 EXPLORATIEF ZOEKEN EN ONDERZOEK

Exploratief onderzoek wordt door White en Roth definieert als een proces dat wordt gekarakteriseerd door een onduidelijke informatiebehoefte aan het begin van het zoekproces (1). Op basis van beschikbare informatie bewegen onderzoekers zich door een informatieruimte en ontdekken ze relevante bronnen (White en Roth 1). Het omvat de studie naar dichte of rijke informatiedomeinen aan de hand van drie voornaamste

facetten: “Lookup”, “Learn” en “Investigate” (Nolan 38; Marchionini ‘Exploratory Search’ 42). Exploratief onderzoek kan worden gezien als een preliminaire onderzoeksfase waarin de onderzoeker zichzelf bekend maakt met ideeën en onderzoeksobjecten (Marchionini ‘Exploratory Search’ 43). Zoals in Kuhlthau’s model beleven onderzoekers in deze fase verschillende gradaties van onzekerheid, die verdwijnen als het onderzoeksproces zich voortzet van exploratief browsen naar gefocust zoeken (White en Roth 1).

Exploratief onderzoek wordt gebruikt om reeds onbepaalde subjecten te onderzoeken, het kan de onderzoeker op ideeën brengen en geeft hem een beeld van de beschikbare informatie. Onderzoekers maken vaak gebruik van exploratief onderzoek wanneer ze bezig zijn met het creëren van een corpus van een onderwerp waar ze geïnteresseerd in zijn (Marchionini ‘Exploratory Search’ 45). Het doen van exploratief onderzoek is ook van belang om te kunnen bepalen of een bepaalde collectie geschikt is om een corpus mee op te bouwen.

Marchionini en White en Roth hebben het exploratieve onderzoeksproces in kaart gebracht en leggen de nadruk op de iteratieve en non-lineaire dynamiek tussen verschillende cognitieve strategieën en hiermee geassocieerde zoektactieken (Marchionini ‘Exploratory Search’ 42; White en Roth 1). In Figuur 2 zijn drie verschillende zoekactiviteiten uitgebeeld door Marchionini waarin hij benadrukt dat exploratief zoeken voornamelijk ‘leer-en onderzoeksactiviteiten’ omvat (Garfield 117; Marchionini ‘Exploratory Search’ 42). Marchionini betoogd dat het opzoeken van informatie iets is wat gebruikt wordt door onderzoekers voor het checken van feiten en bekende objecten (‘Exploratory Search’ 45).

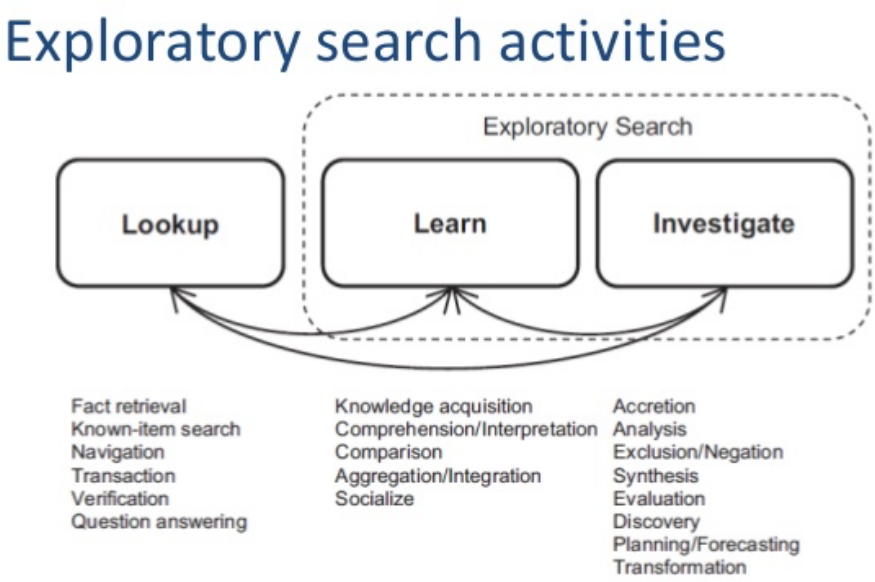
Het model benadrukt de interactie tussen verschillende strategieën en hun geassocieerde tactieken (Crédeville en Forest 168). De leerstrategie doet denken aan de eerste drie fases van Kuhlthau’s model en de eerste fase van Vakkari’s model van het zoeken naar informatie. Het “leren” in het model van Marchionini leidt uiteindelijk tot het verminderen van cognitieve en affectieve onzekerheid door het vermeerderen van de kennis van de onderzoeker.

De relatie tussen deze verschillende facetten zijn niet lineair of hiërarchisch (Marchionini 42; White en Roth 1; Nolan 38). Onderzoekers kunnen meerdere soorten zoekactiviteiten tegelijkertijd inzetten. Nolan geeft ook aan dat sommige van deze activiteiten in andere kunnen zijn ingebed: “behaviors within the “Investigate” activity

may precede those from 'Learn' or 'Lookup', while activities under "Investigate" may be paired with one or more from 'Learn'" (Nolan 38).

Juist deze combinatie en recombinatie van de representatierelaties stelt ons in staat om tot nieuwe ontdekkingen en inzichten te komen (Nolan 38). Deze relaties hebben het potentieel om onderzoekers hun 'eureka moment' te bieden (38). Het doel van informatie-architecten is om gebruikers te ondersteunen in hun vermogen om relaties te construeren tussen gezochte informatie objecten (Gersh et al. 63).

Figuur 2 Marchionini's exploratieve zoekmodel



Marchionini 'Exploratory search: From Finding to understanding '41-46.

2.3 INTERFACE DESIGN

Huurdeman en Kamps geven aan dat user interfaces een cruciale rol spelen in de interactie met zoeksystemen (4). Daarom worden nieuwe interfaces ontwikkeld exploratief onderzoek die rekening houden met relevante zoektactieken en strategieën (Wilson et al. 2; Marchionini 'Exploratory Search' 45). Het doel van het exploreren van informatie is het verfijnen van vage informatiebehoeften door interactie met informatie-objecten en informatiebronnen dat leidt tot een beter begrip van een probleem (White et al. 2).

De onderzoeker moet tijdens het zoeken worden ondersteund om een overzicht te krijgen van de collectie, en bekent te worden gemaakt hiermee. Zoekers kunnen voordeel hebben van interfaces die ondersteuning bieden door bijvoorbeeld resultaten te groeperen of doormiddel van een geleid exploratieproces (Bederson 71; Kules, Shneiderman, en Plaisant 5; White, Jose en Ruthven 1113). Volgens Nolan moet in het design van zoekresultaten en browse-interfaces een bepaalde mate van elasticiteit zitten om gebruikers meer controle te geven over de resultaten (38). Dit kan door exploratieve zoeksystemen gebruik te laten maken van de context van de zoekactiviteit en de contextuele informatie die beschikbaar is, bijvoorbeeld door middel van semantische data. Dit helpt om begrip bij de gebruiker te vergroten en onzekerheid te verkleinen over de aard van het probleem en de informatie die gezocht wordt (White, Khules en Bederson 5).

Verschillende studies laten zien dat een belangrijk interface design principe is om gebruikers onmiddellijk zoekresultaten te laten zien nadat ze een query of navigatie-stap hebben gemaakt (Hutchinson, Bederson en Druin 105; Käki 23; Plaisant et al. 523; Hearst 'Search User Interfaces' 8). Dit helpt de onderzoeker begrijpen of ze op het juiste pad zijn of niet, en het geeft de zoeker ook suggesties van gerelateerde woorden die ze eventueel kunnen gebruiken voor een herformulering van hun initiële zoekterm. Veel experimentele systemen maken de fout om hun gebruikers eerst naar een grote hoeveelheid "helpersinformatie" te laten kijken, zoals queryverfijning suggesties of categorieën, voordat ze zoekresultaten kunnen bekijken (Hearst 'Search User Interfaces' 17).

Een belangrijke vorm van feedback in de zoekresultaten is om de termen uit de query in documenten in de resultaten lijst te laten zien (Hearst 'Search User Interfaces' 18). Dit is om te laten zien hoe het verkregen document zich verhoudt tot de concepten die zijn uitgedrukt in de query. Onderzoek heeft uitgewezen dat het visueel benadrukken van zoektermen een nuttige *feature* kan zijn voor zoekinterfaces (Aula 6; Landauer et al. 73; Lesk 250; Marchionini "Information Seeking in Electronic Environments" 124).

Daarnaast moeten interfaces een divers scala aan zoekstrategieën ondersteunen. zoals het herinneren en herkennen van informatie, *facet*-gebaseerd zoeken en domeinselectie. Onderzoek heeft uitgewezen dat het waarschijnlijk is dat mensen opnieuw informatie bekijken en opnieuw query's gebruiken die ze al een keer eerder

hebben ingetypt (Jones, Dumais en Bruce 400; Milic-Frayling et al. 63). Het mogelijk maken voor gebruikers om te zoeken in recentelijk bekeken informatie kan de productiviteit van de gebruiker verbeteren (Dumais et al. 72). Dit kan bijvoorbeeld door de browsegeschiedenis van de gebruiker bij te houden. (Hearst 'Search User Interfaces' 19). Ook zouden exploratieve zoeksystemen een "workspace" moeten bevatten die een spectrum aan activiteiten zoals notities maken en het ordenen van collecties ondersteund. (Hearst 'Search User Interfaces' 45).

2.3 IMPACT VAN INFORMATIE ZOEKMODELLEN OP INFORMATIEZOEKSYSTEMEN

Het ISP model van Kuhlthau heeft aanzienlijke invloed gehad op bibliotheek- en informatiediensten (Huurdeman en Kamps 3; Kuhlthau 'Accommodating the user's information search process' 12). Haar model heeft echter weinig invloed op het ontwerp van informatiezoeksystemen. Zonder duidelijke richtlijnen te geven op hoe ze moeten worden geïmplementeerd biedt Kuhlthau enige concepten die gebruikt kunnen worden in het ontwerp van informatie zoek systemen. Hieronder vallen het *process* concept en het *uncertainty* principe, de relatie tussen *uniqueness* en *reductancy*, de *mood* of houding van een individueel in het *process*, de evaluerende *interest* en *complexity* van een gebruiker, en het concept van *enough* in het oplossen van een probleem (Huurdeman en Kamps 3).

Daarnaast adviseert Kuhlthau om gebruikers in het begin niet te bedelven onder informatie. Nieuwe *tools* bieden vaak toegang tot een groot aantal bronnen, dit kan verwarrend zijn voor onderzoekers. Ze geeft aan dat een paar "well-chosen interdictory pieces" misschien beter werken in de eerste oriënterende fases van onderzoek ('Seeking meaning' 205).

Vakkari sluit hierbij aan en benadrukt dat in eerdere fases van een taak gebruikers meer ondersteuning nodig hebben, aangezien ze een ongestructureerd mentaal model hebben ('Relevance and contributing information types of searched documents in task performance' 3). In deze fase zijn gebruikers hun raamwerk aan kennis aan het opbouwen van een onderwerp, dit is nodig voordat een focus geformuleerd kan worden (Cole 1218).

De modellen van Kuhlthau en Vakkari beschrijven uiterst nauwkeurig de onderzoeksfasen; ze bieden echter minder handvatten om daadwerkelijk een system te implementeren die de verschillende fases ondersteund. Aangezien hun model het informatiezoekproces op een macro niveau beschrijft kan het moeilijk zijn om specifieke systeem en interface *features* te ontwerpen die worden geleid door de implicaties van modellen op een micro niveau (Wilson 252 ;Huurdeman en Kamps 4).

Onderzoek gedaan naar zoeksystemen op het gebied van interface design levert een aantal 'globale' maar ook expliciete oplossingen voor problemen tijdens de exploratieve onderzoeksfase die ook daadwerkelijk geïmplementeerd kunnen worden.

3. TECHNOLOGIEËN DIE EXPLORATIEF ONDERZOEK ONDERSTEUNEN

Traditionele vormen van zoeksystemen in erfgoedcollecties gaan ervan uit dat een selectie van items tot stand dient te komen die overeenkomt met de ingevoerde query (De Boer et al. 2). Het formuleren van een query, en in het specifiek het herformuleren van een query kan worden begrepen als een van ingewikkeldste taken die gebruikers van interactieve informatiesystemen moeten overwinnen (Belkin et al. 404). Ook Kamps geeft aan dat het formuleren van een expliciete query onpraktisch kan zijn in een exploratieve onderzoeksfase, onder anderen omdat het grote inspanning vergt van de onderzoeker (2). Waitelonis en Sack beamen dit. Volgens hen is zoeken gebaseerd op kernwoorden alleen nuttig als de zoeker exact weet waarnaar zij opzoek is (645). Ze beschrijven de toegevoegde waarde van semantisch zoeken en *linked data* voor exploratieve doeleinden.

Methoden die regelmatig worden ingezet om gebruikers te ondersteunen in exploratief onderzoek zijn het suggereren van query termen, *relevance feedback* (Jin, Sloan en Wang 665) of *interactive user modeling* (Ruotsalo et al. 2). Het geven van snelle toegang tot zoekresultaten door deze te presenteren in een *faceted* browser of het clusteren van zoekresultaten (Yee et al. 401; Hearts en Pendersen 76; Kangasrääsiö 1). In dit hoofdstuk wordt ingegaan op het gebruik van events en narratieven om data te modelleren.

3.1 HET SEMANTISCHE WEB EN *LINKED DATA*

Het semantische web werd door Tim Berners-Lee omschreven als "A new form of Web Content that is meaningful to computers" (Berners-Lee). Het semantische web werd geïntroduceerd als een extensie van het huidige web waarin informatie een goed gedefinieerde betekenis heeft en er voor zorgt dat computers en mensen beter samen kunnen werken (Berners-Lee, Hendler en Lassila 34). Het semantische web gaat over het creëren van betekenisvolle links tussen heterogene data bronnen, dit wordt ook wel *linked data* genoemd. Dit stelt personen en machines in staat om een web aan data te exploreren (Berners-Lee, Hendler en Lassila 35; Waitelonis en Sack 684). Onderlinge verbanden maken het mogelijk om te navigeren van een bron naar andere gerelateerde bronnen uit andere informatiebronnen en hier kennis van op te doen (Waitelonis en Sack 684).

Het Semantische Web is gebaseerd op een *Resource Description Framework* (RDF), dit is een standaard model voor data-uitwisseling op het web. RDF heeft functies die het mogelijk maken data samen te voegen als de onderliggende schema's niet gelijk zijn. Het ondersteund de evolutie van schema's door de tijd heen zonder dat het noodzakelijk is dat al de data consumenten zich moeten aanpassen (W3C). RDF bereidt de gelinkte structuur van het web uit en gebruikt *Uniform Resource Identifier's* (URIs)[8] om relaties tussen objecten te benoemen en om de objecten zelf te identificeren (Waitelonis en Sack 648). RDF wordt dus gebruikt om uitspraken te doen over bronnen op het web. Dit gebeurt over het algemeen in de vorm van een driedelige subject-predicaat-object-structuur[9], ook wel een *triples* genoemd (W3C). *Triples* vormen een geleiden gelabelde graaf waarvan de randen een benoemde link representeren tussen twee bronnen, gerepresenteerd door *nodes* (Swick).

Vergleken met andere manieren van data structureren op het web, biedt *linked data* een enkel gestandaardiseerd toegangsmechanisme. Dit maakt dat de data en de interface los van elkaar staan en gemakkelijk kan worden hergebruikt (Bizer et al. 1265).

3.2 IMPLICATIES LINKED DATA VOOR EXPLORATIEVE ZOEKBROWSERS

De bovengenoemde innovaties op het gebied van *linked data* maken het mogelijk om interactieve toegang te bieden tot heterogene erfgoedcollecties, zowel door directe query zoekacties, alsmede door gestructureerde links tussen cultureel erfgoed objecten, gerelateerde events, personen plaatsen en concepten (De Boer et al. 2). Semantisch zoeken maakt het mogelijk om zoeken gebaseerd op kernwoorden te verbeteren door rekening te houden met de daadwerkelijke *content* van de informatie en zijn semantiek. Door middel van semantische annotatie kunnen informatiebronnen aan elkaar worden gerelateerd. Zo kunnen verborgen en impliciet bestaande relaties kunnen expliciet worden gemaakt (Waitelonis en Sack 646). In plaats van een enkele zoekterm te

[8] URI is een element van een internetprotocol, het is een eenmalige benaming van een bron om deze te kunnen identificeren in het semantische web (W3C URI).

[9] Het subject is in essentie de bron die beschreven wordt. Het predicaat is het kenmerk of aspect van de bron die beschreven wordt. Het object is de waarde van dat kenmerk.

gebruiken kan gebruik gemaakt worden van alle eigenschappen van informatiebronnen en hun relaties onderling om zo geleide exploratie van zoekruimtes mogelijk te maken en hiermee ook de mogelijkheid voor heldere ontdekkingen. Gebruikers worden hebben zo de mogelijkheid zich door de zoekruimte navigeren en te exploreren langs geleide routes (Waitelonis en Sack 646). Een eventuele uitwerking hiervan zou kunnen zijn het presenteren of visualiseren van collectieobjecten in een narratieve structuur.

3.3 EVENTS EN NARRATIEVEN

Van den Akker geeft aan dat "the main building blocks in the interpretation process of history and collection objects are historical events and the relations between these events and the objects" ('Digital Hermeneutics' 1). De culturele en historische betekenis van een object wordt volgens Van den Akker bepaald door zijn relatie tot andere objecten, *events* en zijn plaats in een overkoepelend verhaal ('From Information Delivery to Interpretation Support' 433; 'Digital Hermeneutics' 1). *Events* hebben van zichzelf narratieve elementen. Door *events* aan het object toe te voegen krijgt het object betekenis en wordt het mogelijk het te interpreteren. Een enkel historisch object geeft maar een deel van de gehele historische context van het object, wat maakt dat narratieven belangrijk zijn in het interpretatieproces (Van den Akker 'Digital Hermeneutics' 1).

Narratieven gaan over de manier waarop verhalen verteld of gepresenteerd worden. Een narratief is een verhaal, een aaneenschakeling van chronologisch gestructureerde *events* die betrekking hebben entiteiten, verteld door een 'narratief discours' (Abbott 3; Alleyne 2). Hieruit volgt dat een verhaal door verschillende narratieve discours kan worden weergegeven (Alleyne 2). Er wordt onderscheid gemaakt tussen de *content* van het verhaal en de manier waarop het wordt verteld. Verhalen hebben vaak interacterende karakters, ze kunnen ook een bepaalde reeks aan feiten en observaties aan elkaar verbinden door een terugkomend thema of argument.

Naast dat *events*, zoals Van den Akker aangeeft, belangrijke componenten zijn in het interpretatieproces van de gebruiker, zijn het ook bouwstenen voor een narratief. In een narratieve structuur zitten chronologische relaties tussen verschillende *events* met daarin de suggestie van een causaal verband.

In deze reeks gebeurtenissen zijn events causaal of op andere wijze met elkaar verbonden. *Events* zijn hierin van belang aangezien verhalen iets vertellen over wat er met iets of iemand is gebeurd. Een verhaal is meer dan een opeenvolging van gebeurtenissen het wordt pas een verhaal als er bijvoorbeeld personages terugkomen of bepaalde relaties worden aangetoond.

3.4 TECHNISCHE IMPLEMENTATIE VAN EVENTS EN NARRATIEVEN

De interface van DIVE kan worden gezien als een visualisatie of presentatie van de onderliggende graaf. Doordat de interface zowel horizontaal als verticaal navigeerbaar is, is het mogelijk voor gebruikers zich te bewegen over de graaf van het onderliggende *linked data* netwerk van historische objecten, *events*, plaatsen, personen en concepten die hierin aanwezig zijn.

Zowel de tekstuele omschrijvingen als de descriptieve *meta data* van de objecten in DIVE zijn opgehaald en geconverteerd naar RDF. Uit de tekstuele omschrijvingen worden *events*, plaatsen, personen en concepten gehaald[10] (De Boer et al. 3). In de tweede fase wordt gebruik gemaakt van het *crowdsourcing* platform CrowdThruth[11] dat de resultaten van *Natural Language Processing* verfijnt.

De resultaten van de verschillende *tools* en de *crowdsourcing* opdracht worden geconsolideerd naar RDF. De data is gemodelleerd met gebruik van het *Simple Event Model*[12] (SEM). SEM wordt gebruikt voor het modeleren van *events* gebaseerd op vier elementen: actor, locatie, periode en type (Hage et al. 'The Simple Event Model

---

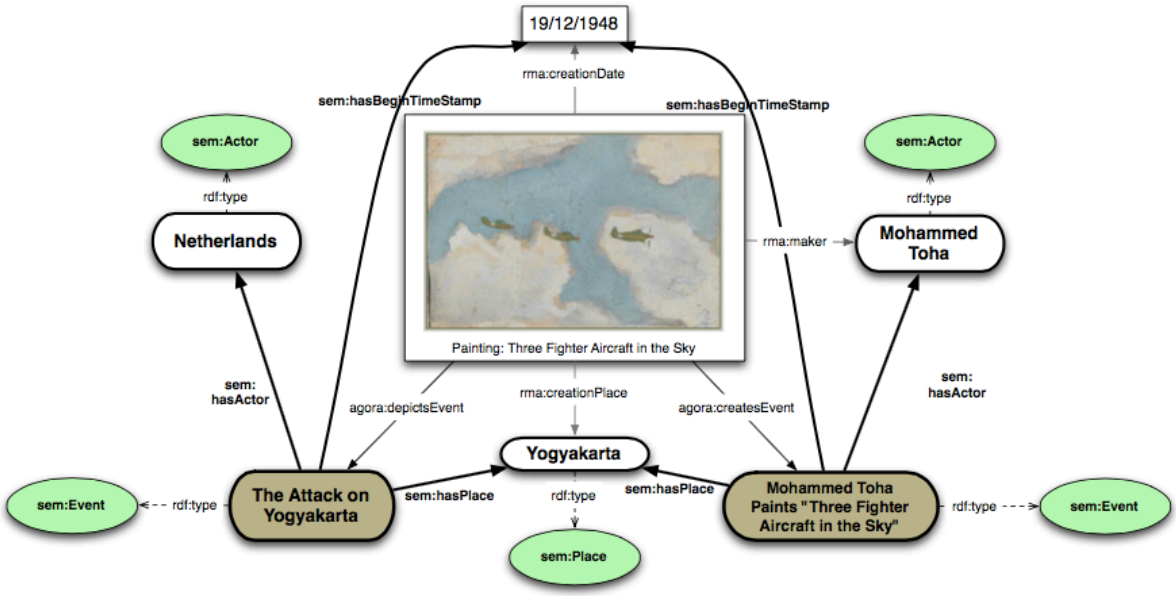
[10] Hiervoor worden een aantal methoden gebruikt waaronder *Named Entity Reconition* (NER) en Event extractie *tools* voor Nederlandse tekst waaronder de XTas en Opener Toolkit.

[11] http:/crowdthruth.org/

[12] SEM is gecreëerd om *events* te modelleren in verschillende subject domeinen. Het is ontwikkeld om aan de ene kant gebruik te maken van bestaande vocabulaire en data bronnen die *events* omschrijven en aan de andere kant om gebruikers in staat te stellen bruikbare feiten af te leiden (Hage et al. 'Design and use of the Simple Event Model (SEM) 128). *Events* beschrijven volgen SEM gebeurtenissen. Het design van SEM stelt het in staat verschillende standpunten in te nemen (Semantic web).

Ontology'). *Events* kunnen in twee verschillende soorten relaties betrokken zijn: *object-event* relaties en *event-event* relaties (Van den Akker 'Digital Hermeneutics' 2). *Object-event* relaties zijn gedefinieerd alszijnde relaties tussen een object en een *event* die betekenisvol kan zijn vanuit een historisch perspectief. Er worden drie types object-event relaties onderscheiden die de *event*-dimensies van een object definiëren. (Van den Akker 'Digital Hermeneutics' 2). De eerste twee typen worden gegeven door de historische context van een object. Deze hebben betrekking op de *content* van een object. Een *object-event* relatie kan een *event* presenteren of definieert een functie in een gebeurtenis. Daarnaast is er nog een derde type *object-event* relatie dat inzicht geeft in de geschiedenis van het object zelf, bijvoorbeeld met betrekking tot de ontstaansgeschiedenis of de herkomst. (Van den Akker 'Digital Hermeneutics' 3).

Figuur 3 Event dimensies van een object



Van den Akker et al. 'Digital Hermeneutics' 3

In figuur 3 is een voorbeeld te zien van een object en haar *event* dimensies. Ook zijn onderlinge event relaties te zien. Het object is een schilderij gemaakt door Mohammed Toha waarop de aanval op Yogyakarta is afgebeeld. De database van de collectie geeft aan dat het object verbonden is met het *event* "Mohammed Toha Paints "Three Fighter Aircraft in the Sky"". Het object heeft een plaats "Yogyakarta" en een

actor “Mohammed Toha”. Het object is gecreëerd op “19 December 1948”. De *event* informatie die gerelateerd is aan het object, verrijkt het originele object. Volgens Van den Akker et al. leiden niet alle *event* dimensies tot dezelfde interpretatie van een object (‘Digital Hermeneutics’ 3). Volgens hem is het “Attack on Yogyakarta” *event* bruikbaarder dan het “Mohammed Toha Paints” *event* voor een onderzoek dat zich richt op bijvoorbeeld oorlog (‘Digital Hermeneutics’ 3). Daarom worden er ook *event-event* relaties gedefinieerd die betekenisvolle relaties tussen *events* vaststellen. *Event-event* relaties worden gedefinieerd door drie elementen: topologie (dezelfde plaats), concept (hetzelfde type) en biografie (dezelfde acteur) (Van den Akker ‘Digital Hermeneutics’ 3). Deze relaties ondersteunen en sturen het interpretatieproces van de gebruiker (Van den Akker ‘Digital Hermeneutics’ 3). Deze *event-event* relaties zijn een goede basis voor het creëren van historische narratieven.

4. VISUALISATIES IN EEN EXPLORATIEVE ONDERZOEKSFASE

Het visueel exploreren van data stelt gebruikers in staat patronen te observeren. Volgens Crédeville en Forest is het visualiseren van zoekresultaten behulpzaam in het informatiezoekproces. (167). Liu et al. analyseren hoe cluster gebaseerde *Exploratory visualization systems* (EVSs) gebruikers ondersteunen in een variatie van verschillende zoekactiviteiten (1). Volgens hen maximaliseren *Exploratory visualization systems* (EVSs) de informatiewinning van gebruikers (Liu et al. 1). Daarentegen geven Granitzer et al. en Hornbæk en Frøkjær aan dat informatievisualisatie-interfaces die documenten als punten of iconen in een twee dimensionale ruimte laten zien, zorgen voor een lage gebruiksvriendelijkheid aangezien gebruikers de tekst van de titels en documenten niet kunnen zien (127; 180).

Informatievisualisaties kunnen worden begrepen alszijnde interacties die grafische waarneming ondersteunen (Bederson 71). Visualisaties staan ten dienst van het analytisch vermogen van de gebruiker (Card, Mackinlay en Shneiderman 10; Crédeville en Forest 168). Volgens Crédeville en Forest lijken deze cognitieve eigenschappen de strategieën en tactieken te ondersteunen die nodig zijn bij het doen van exploratief onderzoek (174).

Een voorbeeld van een *tool* die gebruikt maakt van visualisaties voor het exploreren van data is *Carot2Search.* De zoekbrowser maakt zowel gebruikt van tekstuele als visuele presentaties in de interface. De *tool* clustert termen uit zoekresultaten visueel in een *donut chart* (Crédeville en Forest 173). Het bezwaar van Granitzer et al. en Hornbæk en Frøkjær is hier niet aan de orde aangezien de visualisatie de daadwerkelijke objecten niet vervangt maar een extra analysedimensie biedt. De *Intent Radar* interface ondersteund gebruikers in het ontdekken van de informatieruimte doormiddel van associaties van de geuite informatiebehoefte (Ruotsalo et al. 2; Glowacha et al. 117). Een veelgebruikte vorm van informatievisualisering is het plotten van objecten op een kaart of tijdlijn. Zoals bijvoorbeeld *Historypin*[13] en het project *"A History of the World"*[14].

Volgens Crédeville en Forest is het over het algemeen onduidelijk hoe informatievisualisaties precies moeten worden toegepast in exploratieve zoeksystemen

[13] https://www.historypin.org
[14] http://www.bbc.co.uk/ahistoryoftheworld/explorerflash/

(168). Daarnaast is er volgens Lui et al. een tekort aan een algemeen beoordelingscriterium voor deze systemen (1). Een uitdaging in dit gebied is volgens Lui et al. het begrijpen van hoe visualisaties in zoeksystemen het cognitieve leerproces ondersteunen.

4.1 NARRATIEVE PRESENTATIES

Verhalen gebaseerd op data verschillen op belangrijke punten met traditionele manieren van verhalen vertellen. Segel en Heer stellen dat in de meeste gevallen verhalen in tekst en film een reeks aan *events* in een strak gecontroleerd verloop presenteren (1). Datavisualisaties worden geprezen om het vermogen om verhalen in data bloot te leggen. Gevisualiseerde data kan in een lineaire reeks van vaststaande *events* georganiseerd worden op een manier gelijksoortige met traditionele verhalen. Maar daarnaast kunnen ze ook interactief zijn, uitnodigen voor verificatie, nieuwe vragen oproepen en voor alternatieve verklaringen zorgen (Segel en Heer 1). Aan informatietechnologieën kan worden deelgenomen, ze zijn interactief, dynamisch, collaboratief en maken hiermee communicatie mogelijk. Volgens Van den Akker is "de dialoog" dan ook het onderliggende concept van informatietechnologieën ('History as Dialogue' 103).

Narrativiteit en interactiviteit staan in een bepaalde zin tegenover elkaar (Alleyne 159). Alleyne stelt toch voor interactiviteit toe te voegen aan een narratieve onderzoeksverslag of presentatie. Hij exploreert hoe interactieve narratieven kunnen bijdragen aan de presentatie van onderzoeksresultaten (Alleyne 160). Door *hyperlinks* en *features* van een interactieve *mind map* met elkaar te verbinden in een kaartstructuur kunnen we het makkelijker maken voor lezers om cross-referenties te volgen en onze bevindingen te verkennen (Alleyne 160). Volgens Alleyne bieden interactieve web-gebaseerde media een oplossing voor het presenteren van tekstuele en audiovisuele data waarin het onderzoek is gefundeerd. Door een narratieve vertelling op een interactieve manier te presenteren wordt de narrativiteit van het geheel gereduceerd. Het gepresenteerde verhaal wordt minder lineair door de interactieve *features*. De interactieve presentatie van de bronnen van het onderzoeksverslag gaat ten koste van de narratieve structuur. Volgens Alleyne is de oplossing hiervoor om het tekstuele onderzoek beschikbaar te stellen naast het interactieve verslag. Alleyne betoogt dat het

gehele onderzoek er in termen van communicatie dan alsnog op vooruit gaat (Alleyne 160).

4.2 DIVE INTERFACE

De huidige interface van DIVE kan worden gezien als een presentatie van de onderliggende *linked data* graaf. In de interface zijn een aantal visualisatie-elementen te zien, bijvoorbeeld kleurcodering van de verschillende type entiteiten. De interface van DIVE presenteert data van vier collecties in zowel een horizontale als verticale interface. In figuur 4 is een screenshot te zien van de huidige interface van DIVE. DIVE is wat Bron, van Gorp en de Rijke een *aggregated* zoekinterface zouden noemen (1535). De interface van DIVE biedt gebruikers een overzicht van objecten uit verschillende collecties. Bron, van Gorp en de Rijke onderscheiden twee verschillende type zoekinterfaces, een *tabbed*[15] en een *blended* interface (1535). Een *tabbed* interface geeft gebruikers toegang tot verschillende bronnen in aparte tabs, een *blended* interface combineert meerdere bronnen in een enkele pagina (Mellucci en Baeza-Yates 116). Dit laatste is ook bij DIVE het geval.

De interface is onderverdeeld in een aantal componenten, waaronder een zoekbalk, zoektermen suggesties, objectpresentaties. Ook is er een index te zien met iconen en kleuren die de verschillende typen en het aantal gerepresenteerde entiteiten weergeven. Het is mogelijk om te filteren op basis van type entiteit. De entiteiten indien mogelijk gesorteerd op datum. Het type van een entiteiten is zichtbaar gemaakt door middel van een kleurenbalk. De kleur van het vastgestelde type (*event*, concept, plaats, persoon, of mediaobject) wordt afgebeeld als een entiteit geen afbeelding heeft. De relatie tussen deze entiteiten is echter niet duidelijk weergegeven.

Door met de muis te scrollen is het mogelijk op objecten in te zoomen. Eenmaal ingezoomd krijgt de gebruiker meer informatie over het selecteerde *event*, concept, plaats, persoon, of mediaobject. Op het moment is er weinig tot geen omschrijving te vinden van de verschillende entiteiten in de collectie. Nadat er een entiteit geselecteerd is, wordt in de horizontale lijn hieronder een nieuwe rij aan entiteiten gegenereerd die in verband staan met het geselecteerde object. Dit proces kan oneindig worden herhaald.

---

[15] Mellucci en Baeza duiden dit aan als *non-blended* interface (117).

DIVE poogt exploratief onderzoek te ondersteunen. Deze ondersteuning is op het moment niet ideaal. In een exploratieve onderzoeksfase van moeten onderzoekers ondersteund worden om een overzicht te krijgen van de collectie. Het is op het moment onduidelijk welke erfgoedcollecties aanwezig zijn in de DIVE interface. Het is cruciaal voor wetenschappers om kritisch te kunnen reflecteren op hun bronnen. Ook is het onduidelijk waarom de huidige collecties aanwezig in DIVE samen zijn gevoegd en wat hiervan de toegevoegde waarde is voor de onderzoeker.

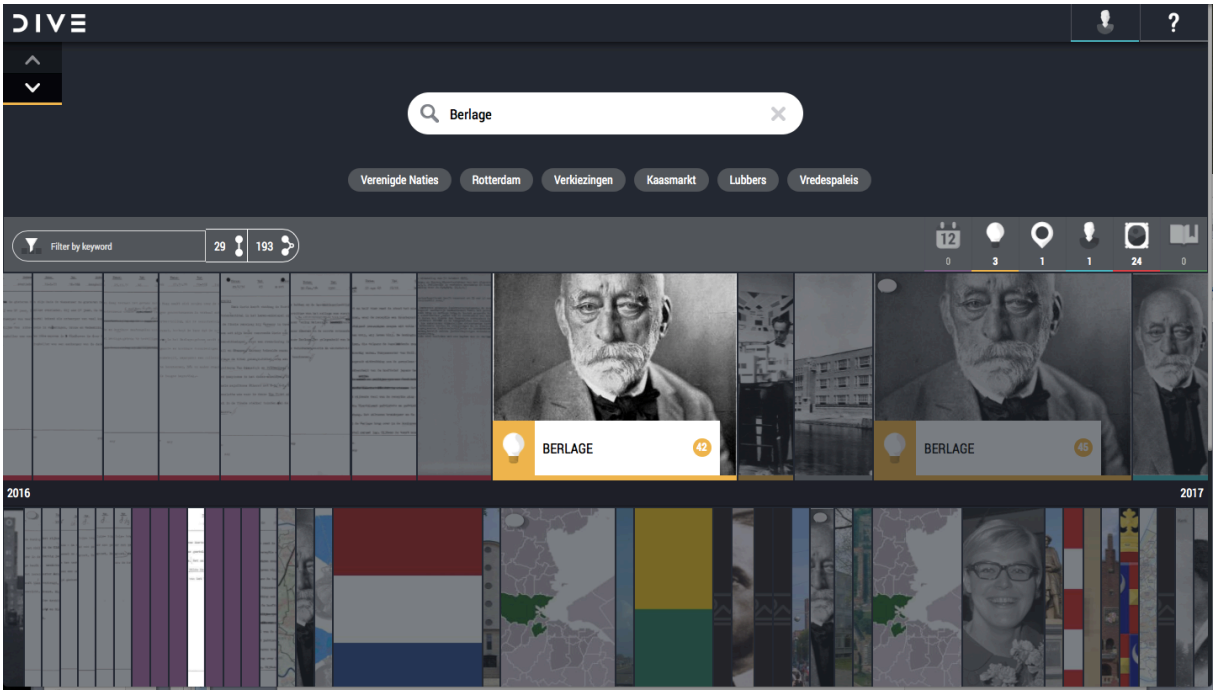
In een exploratieve onderzoeksfase is het belangrijk dat onderzoekers worden ondersteund door ze bijvoorbeeld verschillende voorbeelden en mogelijkheden te laten zien. De zoekbrowser biedt een aantal voorbeelden van zoekopdrachten die zouden kunnen worden ingevuld. Er is echter geen mogelijkheid om andermans zoek of *browse* geschiedenis te bekijken.

Het is op het moment niet mogelijk om een collectie aan objecten te verzamelen of zoekresultaten op te slaan. Dit kan behulpzaam zijn zodat links tussen entiteiten uit de collectie kunnen worden geanalyseerd. In deze nieuwe onderzoeksfase moet de onderzoeker ondersteund worden in de vorm van visualisaties van zijn zoekresultaten.

4.3 Implementatie en presentatie van narratieven in DIVE

DIVE is een *event*-gebaseerde browser, dit geeft een goede aanleiding voor het onderzoeker van presentaties van narratieven in de interface aangezien een narratieve structuur relaties legt tussen *events*. Deze narratieven zouden zowel autogegenereerd kunnen worden of geconstrueerd door gebruikers. Autogegenereerd narratieven, zouden kunnen worden gevormd door een reeks aan *event-event* relaties uit te lichten. Op zich vormen de bestaande *event-event* relaties in DIVE een goede basis voor het creëren van historische narratieven. Van den Akker noemt dit *proto-narratives* ('Digital Hermeneutics' 3). Deze relaties worden bijvoorbeeld gegenereerd als de *events* dezelfde plaats hebben, conceptueel overeenkomen of dezelfde auteur hebben als andere *events*. Er zijn echter nog een groot aantal andere elementen in *event* relaties die samen een narratieve structuur zouden kunnen creëren, maar die niet uitgelicht kunnen worden door een gebrek aan metadata. Dit beperkt het soort en het aantal mogelijke autogegenereerde *proto-narratives.*

Figuur 4 DIVE interface



Hoe passen de *proto-narratives* van Van den Akker in de klassieke narratologie? Zoals al eerder werd aangegeven zijn in de narratologie twee verschillende niveaus te onderscheiden. Allereerst is er de geschiedenis of verhaal en daarnaast het discours. Het verhaal is een chronologische aaneenschakeling van gebeurtenissen (*events*) die betrekking hebben op entiteiten (Abbott 3), het verhaal wordt verteld of gepresenteerd door een narratief discours (soms ook aangeduid als plot): de "story as narrated" (Abbott 3). In het verhaal staan de gebeurtenissen centraal en gaat het om wat er wordt verteld. In de vertelling van het verhaal staat de manier waarop het verhaal wordt gepresenteerd centraal. Een vertelling is ook datgene wat een geschiedenis presenteert. De ordening van gebeurtenissen wordt ook wel het plot genoemd. Alleyne vergelijkt het plot met een graaf die een bepaalde schikking of presentatie van fundamentele data uit een verhaal geeft. Het plot in een narratief is net zoals een plot in een datavisualisatie ,het heeft betrekking op de organisatie van fundamentele data (Alleyne 3).

In het geval van de autogegenereerde narratieven is het onderliggende algoritme verantwoordelijk voor reeks aan gebeurtenissen en hoe deze wordt gepresenteerd; het narratieve discours. Vragen die hierbij opkomen zijn: kan het algoritme de rol van *narrator* op zich nemen, en hoe bepaalt het algoritme de volgorde waarop het een reeks aan gebeurtenissen presenteert? Autogegenereerde narratieven kunnen functioneren

als clusters van informatie. Het groeperen van informatie ondersteunt van het interpretatieproces van de onderzoeker (Bederson 71; Kules, Shneiderman, en Plaisant 5; White, Jose en Ruthven 1113).

Onderzoekers de mogelijkheid geven zelf narratieven te construeren aan de hand van collectie-entiteiten kan bijdragen aan de interpretatie en contextualisatie van andere onderzoekers. Door deze narratieven vervolgens in de DIVE interface te presenteren. Hiervoor zijn er wel een aantal extra *features* nodig in de DIVE interface. Voor het construeren van een narratief is het allereerst nodig om objecten in een 'collectie' op te kunnen slaan, dit draagt ook bij aan het creëren van een corpus voor wetenschappelijk onderzoek, wat een belangrijk onderdeel is van exploratief onderzoek (Marchionini 'Exploratory Search' 45). Daarbij is uit onderzoek gebleken dat onderzoekers vaak informatie opnieuw opzoeken (Jones, Dumais en Bruce 400; Milic-Frayling et al. 63). Daarom kan het bijdragen aan de productiviteit van de gebruiker door haar de mogelijk te geven objecten in een 'collectie' op te slaan, of de aaneenschakeling van navigatieoperaties beschikbaar te maken (Dumais et al. 72; Hearst 'Search User Interfaces' 19).

Gebruikers hebben daarnaast een soort 'workspace' nodig om de collectie te *editen*/hergroeperen, waarin zij ook eventueel een narratief kan construeren (Hearst 'Search User Interfaces' 45). Het is van belang dat gebruikers links tussen entiteiten in narratieven kunnen specificeren. Dit draagt bij aan het *linked data* netwerk van objecten, events, concepten en plaatsen. In deze *workspace* moet de onderzoeker ondersteund worden om de objecten te selecteren en ondervragen. Dit kan bijvoorbeeld aan de hand van een tijdlijn of kaart.

Om narratieven te kunnen gebruiken voor een exploratieve onderzoeksfase moeten deze al aanwezig zijn in de browser. Dit kan worden opgelost door collectie-eigenaren, wetenschappers te benaderen van deze narratieven te construeren in de DIVE browser. De opgeslagen narratieven kunnen dienen als referentiepunt of aanleiding voor verder onderzoek. Deze kunnen worden gebruikt als duidelijk voorbeeld ter introductie van de DIVE interface. Ook geeft het de onderzoeker enige houvast. Dit is van belang aangezien onderzoekers in de exploratieve fase van onderzoek onzeker zijn en niet goed kunnen aangeven naar wat voor soort informatiebronnen ze precies opzoek zijn. Een louter gekwantificeerde visuele representatie van narratieven of collectieobjecten is niet bevorderlijk voor de

gebruiksvriendelijkheid van een interface in een exploratieve onderzoeksfase (Granitzer et al. 127; Hornbæk en Frøkjær 180). Wel kunnen visualisatie-elementen het exploratieproces ondersteunen. Dit kan bijvoorbeeld door de daadwerkelijke objecten uit te lichten in een narratieve structuur. Volgens Alleyne kan het interactief presenteren van een narratief onderzoek het best samengaan met het tekstuele onderzoek (159). De interactiviteit van digitale *tools* zorgt voor een reductie van de narrativiteit van het onderzoek zijn geheel. De interactieve presentatie van de bronnen van het onderzoeksverslag gaat ten koste van de narratieve structuur van het verhaal. Hij biedt hiervoor als oplossing het tekstuele onderzoek naast het interactieve verslag beschikbaar te stellen. Volgens hem gaat het gehele onderzoek er in termen van communicatie dan op vooruit (Alleyne 160).

5. CONCLUSIE EN VERVOLGONDERZOEK

In het digitale tijdperk wordt online zoeken steeds belangrijker. Informatie is in steeds grotere mate online beschikbaar, dit geldt ook voor collecties van erfgoedinstellingen. Er is een groeiende behoefte aan het integreren van verschillende culturele collecties, en nieuwe manieren om hierin te navigeren, onder anderen voor het doen van wetenschappelijk onderzoek (Brown 4; Evens en Hauttekeete 163). Wetenschappers hebben *tools* nodig om door deze enorme collecties te navigeren. Er zijn veel verschillende factoren waar rekening mee moet worden gehouden bij het ontwerpen van een *tool* die onderzoekers poogt te ondersteunen in het doen van onderzoek.

Er zijn bijvoorbeeld verschillende fases van onderzoek te onderscheiden, in deze fases hebben onderzoekers verschillende informatiebehoeften. Dit heeft implicaties voor het ontwerp van informatiezoeksystemen. De eerste exploratieve fase van onderzoek wordt gekarakteriseerd door een onduidelijke informatiebehoefte aan het begin van het zoekproces (White en Roth 1). In deze fase van onderzoek speelt onzekerheid een grote rol (Kuhlthau 369-70). Ook hebben onderzoekers in deze fase vaak geen duidelijke focus of doel (Vakkari 'A theory of the task-based information retrieval process' 46).

In tegenstelling tot de gangbare theorie over zoekmodellen, levert literatuur over interface design concrete oplossingen voor problemen die onderzoekers ervaren in een preliminaire onderzoeksfase. Interfaces die zoekresultaten groeperen of de zoeker sturen door een geleid exploratieproces hebben bewezen nuttig te zijn (Bederson 71; Kules, Shneiderman, en Plaisant 5; White, Jose en Ruthven 1113). Het gelijk laten zien van zoekresultaten is behulpzaam. Daarbij is een belangrijke vorm van feedback, het laten zien van zoektermen in de resultaten (Hearst 'Search User Interfaces' 18). Zo kan de zoeker zien hoe het verkregen document zich verhoudt tot de ingevoerde query (Aula 6; Landauer et al. 73; Lesk 250; Marchionini "Information Seeking in Electronic Environments" 124).

Daarnaast zijn er ook technische ontwikkelingen die kunnen bijdragen aan het ondersteunen van een onderzoeker in de exploratieve fase van onderzoek. Bijvoorbeeld door zoeksystemen gebruik te laten maken van de context van de zoekactiviteit en de contextuele informatie die beschikbaar is. Innovaties op het gebied van *Linked data* maken het mogelijk om toegang te bieden tot heterogene erfgoedcollecties door middel van gestructureerde links tussen cultureel erfgoed objecten, gerelateerde events,

personen plaatsen en concepten (De Boer et al. 2). Semantisch zoeken maakt het mogelijk om kernwoord gebaseerd zoeken te verbeteren door rekening te houden met de daadwerkelijke *content* van de informatie en zijn semantiek. Door middel van semantische annotatie kunnen informatiebronnen aan elkaar worden gerelateerd, verborgen en impliciet bestaande relaties kunnen expliciet worden gemaakt (Waitelonis en Sack 646). Dit maakt dat gebruikers zich langs geleide routes door de zoekruimte kunnen navigeren (Waitelonis en Sack 646).

Naar aanleiding van onderzoek van Van den Akker is onderzocht of narratieve presentaties bij zouden kunnen dragen aan het interpreteren en contextualiseren van erfgoedobjecten in de zoekbrowser DIVE. *Events* staan centraal in DIVE. Door middel van object-event en event-event relaties te definiëren ontstaat een *linked data* netwerk van objecten, concepten, plaatsen, personen en type. Events zijn ook een kernonderdeel van een narratieve structuur. Het voorgestelde concept van autogegenereerde *proto-narratives* brengt een aantal vragen met zich mee. Ervan uitgaande dat een narratief bestaat uit een verhaal, een aaneenschakeling van gestructureerde gebeurtenissen en een manier waarop het verhaal wordt gepresenteerd of verteld. Het narratief discours doet de vraag opkomen of een algoritme kan functioneren als *narrator.* Daarnaast is het van belang inzicht te hebben in hoe de presentatie van het narratief tot stand is gekomen, hoe en waarom zijn de events in de huidige volgorde gepresenteerd?

Ook is er de mogelijkheid om onderzoekers zelf de mogelijkheid te bieden om narratieven te construeren binnen de DIVE collectie. Hiervoor zou de browser een aantal nieuwe features moeten toevoegen zoals een *workspace,* waarin entiteiten kunnen worden opgeslagen en vervolgens worden gerangeerd in een narratieve structuur. Deze collectie kan bijdragen aan de productiviteit van de onderzoekers aangezien zij vaak bronnen opnieuw bekijken (Jones, Dumais en Bruce; Milic-Frayling et al. 63; Dumais et al. 72). Een overzicht van onderzochte entiteiten of de browsegeschiedenis van een onderzoeker wordt ook als behulpzaam ervaren (Hearst 'Search User Interfaces' 19).

In deze *workspace* zouden ook andere acitiveiten zoals het maken van notities of het bijvoegen van een onderzoek kunnen worden gefaciliteerd (Hearst 'Search User Interfaces' 45). Als de opgeslagen narratieven van onderzoekers worden opgenomen in het *linked data* netwerk kunnen deze in de browser worden gepresenteerd. Dit brengt ook mogelijkheden met zich mee voor het publiceren van bronnen, of het delen van

narratieven onder onderzoeker. Volgens Alleyne doet het presenteren van een narratief onderzoeksverslag op een interactieve manier af aan de narrativiteit van het geheel. Dit kan worden opgelost door een tekstueel verslag te presenteren naast het interactieve verslag. Uiteindelijk gaat het geheel er dan in termen van communicatie op vooruit (Alleyne).

In theorie draagt het visualiseren van informatie bij aan het ontdekken van patroneren en relaties (Crédeville en Forest 174). Deze cognitieve eigenschappen lijken strategieën en tactieken te ondersteunen die van belang zijn in het doen van exploratief onderzoek (Crédeville en Forest 174). Hoe deze informatievisualisaties precies moeten worden toegepast in exploratieve zoeksystemen blijft echter onduidelijk (Crédeville en Forest 168). Daarnaast is er een gebrek aan een algemeen evaluatieframework van deze systemen. De vraag hoe visualisaties in zoeksystemen het cognitieve leerproces ondersteunen vormt nog altijd een uitdaging (Lui et al. 1). In theorie lijkt het implementeren van (autogegenereerde) narratieve presentaties in de DIVE browser bij te dragen aan exploratie van de collectie. Beide zijn echter nog niet genoeg onderzocht in de praktijk.

Vervolgonderzoek kan zich richten op een beter begrip van hoe visualisaties cognitieve leerprocessen ondersteunen. Hierbij moet rekening worden gehouden met de verschillende methoden die onderzoekers gebruiken in het doen van onderzoek. Ook moet rekening gehouden worden met hoe onderzoekers interacteren met het zoeksysteem. Het gebruik van digitale *tools* voor het doen van wetenschappelijk onderzoek verandert zowel de manier waarop onderzoekers onderzoek doen als de vragen die zij stellen en hoe ze deze beantwoorden (van Dijck 1; Borgman 8; Collins en Micheal 138; Michel et al. 1; Toms en O'Brien 102). Voor het ontwikkelen van deze *tools* is een gedetailleerd begrip nodig van de relatie tussen de verschillende databronnen, *tools* en de activiteiten in het onderzoeksproces. Het is bovendien van belang om te kijken naar hoe deze interacties invloed hebben op de onderzoeksvraag. (Bron, van Gorp en de Rijke 1535). Voor de verdere ontwikkeling van narratieve presentaties in DIVE is het van belang om prototypen te ontwikkelen en deze te testen met de desbetreffende doelgroep.

6. REFERENTIES

Abbott, Porter H. *The Cambridge Introduction to Narrative*. Cambridge: Cambridge University Press, 2008.

Alleyne, Brian. *Narrative Networks: Storied Approaches in a Digital Age.* Londen: SAGE Publications, 2014.

Aula, Anne. Enhancing the Readability of Search Result Summaries. *Proceedings of HCI 2004* (2004): 6–10. 10 mei 2016. <https://www.researchgate.net/profile/Anne_Aula/publication/220969416_Query_Formulation_in_Web_Information_Search/links/53da83e60cf2e38c6336bedd.pdf>.

Bederson, Ben B. PhotoMesa: A Zoomable Image Browser Using Quantum Treemaps and Bubblemaps. *Proceedings of the Conference on User Interface and Software Technology* (2001): 71-80. 20 april 2016. <http://www.dtic.mil/dtic/tr/fulltext/u2/a412705.pdf>.

Belkin, Nicholas, Jona Seeger en Gernot Wersig. "Distributed expert problem treatment as a model for information system analysis and design." *Journal of Information Science* 5.5 (1983): 153–167. 10 juni 2016.< http://jis.sagepub.com/content/5/5/153.full.pdf+html>.

Berners-Lee, Tim. "Linked data. World wide web design issues." *W3.* 2006. The World Wide Web Consortium. 20 april 2016. <http://www.w3.org/DesignIssues/LinkedData.html>.

Berners-Lee, Tim, James Hendeler en Ora Lassila. "The Semantic Web: a new form of Web content that is meaningful to computers will unleash a revolution of new possibilities." *Scientific American* 284.5 (mei 2001): 34-43. 20 april 2016. <http://www.scientificamerican.com/article/the-semantic-web/>.

Bizer, Christian, et al. "Linked data on the web." *WWW '08 Proceedings of the 17th international conference on World Wide Web* (2008): 1265-1266. 10 juni 2016. <http://dl.acm.org/citation.cfm?id=1367760>.

Borgman, Christine L. "The Digital Future is Now: A Call to Action for the Humanities" *Digital Humanities Quarterly* 4.1 (2010): 1-22. 5 maart 2016<http://works.bepress.com/borgman/233>.

Bron, Mark, Jasmijn van Gorp en Maarten de Rijke. "Media studies research in the data-driven age: How research questions evolve." *Journal of the Association for*

Information Science and Technology 67.7 (2015): 1535-1554. 10 mei 2016. <http://doi.org/10.1002/asi.23458>.

Brown, Stephen, Simon Coupland en David Croft. "Picture Perfect: Computational methods for matching historical photographic records across different collections." *Proceedings of the Digital Research in the Humanities and Arts conference 2014* (2014): 17-22. 2 juni 2016.<https://www.researchgate.net/publication/285588409_Picture_Perfect_Computational_methods_for_matching_historical_photographic_records_across_different_collections>.

Brusilovsky , Peter, en Mark T Maybury. "From adaptive hypermedia to the adaptive web." *CACM* 45.5 (2002): 30-33. 2 juni 2016.<http://dl.acm.org/citation.cfm?id=506239>.

Capurro, Rafael. "Digital hermeneutics. An outline." *AI & Society* 35.1 (2010): 35-42. 10 maart 2016. <http://www.capurro.de/digitalhermeneutics.html>.

Card, San K, Jock D. Mackinlay en Ben Shneiderman. *Readings in Information Visualization: Us- ing Vision to Think*. San Francisco: Morgan Kaufmann Publishers, 1999.

Checkland, Peter en Sue Holwell. *Information, Systems and Information Systems: Making Sense of the Field*. Chichester, New York: John Wiley, 1998.

Clariah. Common Lab Research Infrastructure for the Arts and Humanities. 12 april 2016. <http://www.clariah.nl/over>.

Cole, Charles. "A Theory of information need for information retrieval that connects information to knowledge." *JASIST* 62.7 (2001): 1216-1231. 15 mei 2017. <http://onlinelibrary.wiley.com/doi/10.1002/asi.21541/epdf>.

Collins, Ellen, en Michael Jubb (2012). "How do researchers in the humanities use information resources?" *LIBER Quarterly 21.2* (2012): 176–187. 10 maart 2016. <https://www.liberquarterly.eu/articles/10.18352/lq.8017/>.

Creswell, John W. *Research design: Qualitative, quantitative, and mixed methods approaches.* Londen: Sage Publications. 2009.

Crédeville, Aline en Dominic Forest. Is Visualization Usable for Displaying Web Search Results in an Exploratory Search Context? <http://link.springer.com/chapter/10.1007/978-3-642-36415-0_11>.

De Boer, Victor, et al. “Dive into the Event-Based browsing of Linked Historical Media.” *Web Semantics: Science, Services and Agents on the World Wide Web* 35.3 (December 2015): 152-158. 5 maart 2016. <http://www.sciencedirect.com/science/article/pii/S1570826815000499>.

De Jong, Franciska, Roeland Ordelman en Stef Scagliola. (2011). “Audio-visual Collections and the User Needs of Scholars in the Humanities: a Case for Co-Development.” *2nd Conference on Supporting Digital Humanities* (2011), 10 april 2016. <http://eprints.eemcs.utwente.nl/20868/ >.

Dervin, Brenda. “An overview of sense-making research: concepts, methods and results to date.” *International Communications Association Annual Meeting*. Dallas, Texas, 1983. 5 mei 2016. <http://faculty.washington.edu/wpratt/MEBI598/Methods/An%20Overview%20of%20Sense-Making%20Research%201983a.htm/>.

Dijck, José van. “CLARIAH: Big Data, Grand Challenges,” Common Lab Research Infrastructure for the Arts and Humanities. 2014. 5 mei 2016. <http://www.clariah.nl/files/pers/Artikel_voor_de_nieuwsmedia_CLARIAH_2014.pdf>.

Dumais, Susan. et al. “Stuff I've Seen: A System For Personal Information Retrieval And Re-Use.” *Proceedings of the 26th Annual International ACM SIGIR Conference on Research and development in information retrieval SIGIR* 3 (2003): 72–79. 2 juni 2016. <http://dl.acm.org/citation.cfm?id=860451>.

Evens, Tom, Laurence Hauttakeete. “Challenges of digital preservation for cultural heritage institutions.” *Journal of Librarianship and Information Science* 43.4 (2011): 157-165. 10 maart 2016. <http://lis.sagepub.com/content/43/3/157.abstract>.

Foster, Allen E. “A nonlinear model of information-seeking behavior.” *Journal of the American Society for Information Science and Technology* 55.3 (februari 2012): 228-237. 1 maart 2016.<http://onlinelibrary.wiley.com/doi/10.1002/asi.10359/full>.

Gall, Medredith D, Walter R, Borg en Joyce P Gall. *Education research: An introduction* New York: Longman, 1996.

Garfield, Eugene. “When is a negative search result positive?” *Essays of an Information Scientist* 1 (augustus 1970): 117– 118. 8 april 2016.<http://www.garfield

.library.upenn.edu/essays/V1p117y1962-73.pdf>.

Gersh, John, et al. “Supporting insight-based information exploration in intelligence analysis.” *Communications of the ACM 49.*4 (april 2006): 63-68. 10 mei 2016. <http://portal.acm.org/citation.cfm?id=1121949.1121984>.

Glowacha, Dorota, et al. ”Directing exploratory search: reinforcement learning from user interactions with keywords.” *IUI’13* (2013): 117-128. 2 juni 2016. <http://dl.acm.org/citation.cfm?id=2449413>.

Granitzer, Michael. et al. “Evaluating a System for Interactive Exploration of Large, Hierarchically Structured Document Repositories.” *Proceedings of the IEEE Symposium on Information Visualization* (2004): 127–133. <http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=1382900&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxpls%2Fabs_all.jsp%3Farnumber%3D1382900>.

Hage, Willem R van, et al. “Design and use of the Simple Event Model (SEM).” *Web Semantics: Science, Services and Agents on the World Wide Web 9* (2011): 128–136 . 20 mei 2016. <http://semanticweb.cs.vu.nl/2009/11/sem/>.

Hage, Willem R. van, et al. *The Simple Event Model Ontology* 20 mei 2016. <http://semanticweb.cs.vu.nl/2009/11/sem/>.

Hearst, Marti A. *Search user interfaces*. Cambridge: Cambridge Univesity Press, 2009.

Hearst, Marti A. “TileBars: Visualization of term distribution information in full text information access.” *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems* (1995): 59-66. 20 mei 2016. <http://dl.acm.org/citation.cfm?id=223912>.

Hearst, Marti A, et al. “Finding the flow in web site search.” *Communications of the ACM*, 45.9 (2002) :42–49. 20 maart 2016. <http://flamenco.berkeley.edu/papers/cacm02.pdf>.

Hearst, Marti A, en Jan O. Pedersen. “Reexamining the cluster hypothesis: Scatter/gather on retrieval results.” In *Proceedings of the 19th Annual ACM International Conference on Research and Development in Information Retrieval* (1996): 76–84. 20 maart 2016. <http://dl.acm.org/citation.cfm?id=243216>.

Hutchinson, Hilary B, Benjamin B. Bederson en Allison Druin. “The Evolution of the International Children's Digital Library Searching and Browsing Interface.” *Interaction Design And Children* IDC '06 Proceedings of the 2006 conference on

Interaction design and children (2006): 105–112. 10 mei 2016. <http://dl.acm.org/citation.cfm?id=1139101>.

Huurdeman, Hugo. C, Jaap Kamps. “From Multistage Information-seeking Models to Multistage Search Systems.” *Proceedings of the 5th Information Interaction in Context Symposium* (2014): 145–154.16 mei 2016. <http://doi.org/10.1145/2637002.2637020 >.

Hornbæk, Kasper, Erik Frøkjær. “Do Thematic Maps Improve Information Retrieval.” *Human-Computer Interaction INTERACT'99 (1999):* 179–186. 26 mei 2016. <http://www.diku.dk/OLD/undervisning/2003f/datVsystem/papers/Interact99.pdf>.

Ingwersen, Peter. “Cognitive perspectives of IR interaction: elements of a cognitive IR theory.” *Journal of Documentation* 51.1 (1996): 3–50. 10 mei 2016. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.116.2558&rep=rep1&type=pdf>.

Ingwersen, Peter, Kalervo Järvelin. ”Information Retrieval in Context - IRiX.” *ACM SIGIR Forum* 39.2 (December 2005): 1-9. 2 juni 2016<http://sigir.org/files/forum/2005D/2005d_sigirforum_ingwersen.pdf>.

Järvelin, Kalervo. “On information, information technology and the development of society: an information science perspective.” *Information technology and information use*. Reds. Peter Ingwersen, Leif Kajberg en Annelise M. Pejtersen, London: Taylor Graham, 1987. 35–55.

Jin, Xiaoran, Marc Sloan en Jun Wang. “Interactive Exploratory Search for Multi Page Search Results.” *WWW '13 Proceedings of the 22nd international conference on World Wide Web* (2013): 655-666. 2 juni 2016. <http://dl.acm.org/citation.cfm?id=2488446>.

Jones, William, Susan Dumais en Harry Bruce. “Once Found, What Then? A Study Of Keeping Behaviors In The Personal Use Of Web Information” *Proceedings Of The American Society For Information Science And Technology* 39.1 (2002): 391–402. 2 mei 2016 . <http://onlinelibrary.wiley.com/doi/10.1002/meet.1450390143/epdf>.

Käki, Mika. “Enhancing Web Search Result Access With Automatic Categorization.” *Interactive Technology* 2 (2005) 2 juni 2016.<https://tampub.uta.fi/bitstream/handle/10024/96850/978-951-44-9776-6.pdf?sequence=1>.

Kamps, Jaap. “Toward a model of interaction for complex search tasks.” ESAIR’11 (oktober 2011): 7-9. 2 juni 2016. <http://humanities.uva.nl/~kamps/publications/2011/kamp:towa11.pdf

Kamps, Jaap et al. “Articulation information needs in XML query languages. ACM Transactions on Information Systems 24.4 (2006): 407-436. 2 juni 2016. <http://dl.acm.org/citation.cfm?id=1185879>.

Kangasrääsiö, Antti, et al. “Interactive Visualization of Search Intent for Exploratory Information Retrieval” *ICML 2014 Workshop on Crowdsourcing and Human Computing* (2014): 1-5. 2 juni 2016. <https://users.ics.aalto.fi/akangasr/publications/ICML2014_preprint.pdf>.

Kules, Bill, Ben Shneiderman“Users Can Change their Web Search Tactics: Design Guidelines for Categorized Overviews.” *Information Processing and Management* 44.2 (2008): 463–484. 2 april 2016. <http://www.sciencedirect.com/science/article/pii/S0306457307001574>.

Kules, Bill, Ben Shneiderman en Catherine Plaisant “Data exploration with paired hierarchical visualizations: Initial designs of PairTrees.” *Proceedings of the National Conference on Digital Government Research* (2003) 3 april 2016. <http://hcil2.cs.umd.edu/trs/2003-09/2003-09.pdf>.

Kuhlthau, Carol C, Jannica Heinström en Ross J. Todd “The ‘information search process’ revisited: is the model still useful?” *Information research* 13.4 (December 2008): 1-11. 1 juni 2016. <http://www.informationr.net/ir/13-4/paper355.html>.

Kuhlthau, Carol C. “Accommodating the user’s information search process: challenges for information retrieval system designers.” *Bulletin of the ASIST*, 25.3 (1999):12-16. 5 april 2016.<http://onlinelibrary.wiley.com/doi/10.1002/bult.115/epdf>.

Kuhlthau, Carol. C. “*Kuhlthau’s information search process*.” *Theories of information behavior*. Reds. Karen E Fisher, Sandra Erdelez en Lynne E. F. McKechnie. New Jersey: Information Today, Inc., 2006. 230-5.

Kuhlthau, Carol C. “Perceptions of the information search process in libraries: a study of changes from high school through college.” *IPM* 24 (1988): 419–427. 2 mei 2016. <http://faculty.washington.edu/harryb/courses/INFO310/Kuhlthau.pdf>.

Kuhlthau, Carol C. *Seeking meaning: A Process Approach to Library and Information Services.* Norwood: Libraries Unlimited, 2003.

Landauer, Thomas, et al. “Enhancing the usability of text through computer delivery and formative evaluation: the superbook project.” *Hypertext: A Psychological Perspective.* Reds. Cliff McKnight, Andrew Dillon en John Richardson. New York: Ellis Horwood, 1993. 71–136. 5 juni 2016.<https://www.ischool.utexas.edu /~adillon/BookSynopses/HaPP/chapter5.html>.

Lesk, Michael. *Practical Digital Libraries; Books, Bytes, & Bucks*. San Francisco: Morgan Kaufmann, 1997.

Liu, Yujie, et al. “Evaluating exploratory visualization systems: A user study on how clustering-based visualization systems support information seeking from large document collections.” *Information Visualization* 12.1 (2013): 25-43. 3 mei 2016. <http://ivi.sagepub.com/content/early/2012/10/04/147387161245 9995>.

Marchionini, Gary, en Anita Komlodi. “Design of interfaces for information seeking.” *Annual review of information science and technology* 33 (1999): 89-130. 3 juni 2016. <https://www.researchgate.net/publication/243776803_Design_of _interface_for_information_seeking>.

Marchionini, Gary. “Exploratory Search: From Finding to Understanding.” *Communication of the ACM*. 49.4. (April 2006): 41-6. 2 juni 2016. <https://www.ischool.utexas.edu/~i385t-sw/readings/Marchionini-2006-Exploratory_Search.pdf>.

Marchionini, Gary. *Information Seeking in Electronic Environments*. Cambridge: Cambridge University Press, 1995.

Melucci, Massimo, Ricardo Baeza-Yates. *Advanced topics in information retrieval*. New York: Springer,2011.<https://books.google.nl/books?hl=nl&lr=&id=yFD0YjXoii YC&oi=fnd&pg=PR5&dq=Melucci,+M.+R.+BaezaYates.+Advanced+topics+in+info rmation+retrieval,+&ots=p5fqltxfnf&sig=3deOLayeREzgRR0S2v1s8FvYCys#v=on epage&q=aggregated&f=false>.

Michel, Jean-Baptiste, et al. “Quantitative analysis of culture using millions of digitized books.” *Sciencexpress* 331.6014 (2011): 176–182. 3 juni 2016.<http://www.librarian.net/wp-content/uploads/science-googlelabs.pdf>.

Milic-Frayling, Natasa et al. “Smartback: supporting users in back navigation.” *Proceedings of the 13th International Conference on World Wide Web* (2004): 63–71, 2004. 3 mei 2016. <http://dl.acm.org/citation.cfm?id=988682>.

Pharo, Nils, Ragnar Nordlie. “Examining the effect of task stage and topic knowledge on searcher interaction with a "digital bookstore".” *IIiX* (2012): 4–11. 2 juni 2016.< http://dl.acm.org/citation.cfm?id=2362730>.

Plaisant, Catherine, et al. “Bringing treasures to the surface: iterative design for the Library of Congress National Digital Library Program.” *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (1997)*: 518–52. 2 mei 2016. <http://dl.acm.org/citation.cfm?id=259009>.

Ruotsalo, Tuukka, et al. “Supporting exploratory search tasks with interactive user modeling.” *Proceedings of the American Society for Information Science and Technology* 50.1 (2013): 1-10. 2 juni 2016.<https://www.asis.org/asist2013/ proceedings/submissions/papers/29paper.pdf.>.

Ruthven, Ian. “Interactive information retrieval.” *Annual review of information science and technology* 42 (2008): 43–91, 2008. 3 mei 2016.<http://onlinelibrary. wiley.com/doi/10.1002/aris.2008.1440420109/epdf>.

Ross, Seamus, Margaret Hedstrom. “Preservation research and sustainable digital libraries.” *International Journal on Digital Libraries* 5.4 (augustus 2005): 317-324. 3 april 2016. <http://link.springer.com/article/10.1007/s00799-004-0099-3#/page-1>.

Shneiderman, Ben, Catherine Plaisant. *Designing the User Interface.* Boston: Person, 2005.

Segel, Edward en Jefferey Heer. “Narrative Visualization: Telling Stories with Data.” *IEEE Trans. Visualization & Comp. Graphics* (2010): 1-10. 2 mei 2016. <http://vis.stanford.edu/files/2010-Narrative-InfoVis.pdf>.

Sem. Semantic web. 1 juni 2016. <http://semanticweb.cs.vu.nl/2009/11/sem/>.

Spink, Amanda, Howard Greisdorf en Judy Bateman. “From Highly Relevant to not Relevant: Examining Different Regions of Relevance.” *IPM* 34 (1998): 599–62. 2 mei 2016.<http://www.sciencedirect.com/science/article/pii/S0306457398000259>.

Toms, Elaine G, Heather L. O’Brien.“Understanding the information and communication technology needs of the e-humanist.” *Journal of Documentation* 64.1 (2008): 102–130. <http://www.emeraldinsight.com/doi/pdfplus/10.1108/00220410810844178>.

Vakkari, Pertti. “A theory of the task-based information retrieval process: a summary and generalization of a longitudinal study.” *Journal of Documentation* 57.1 (2001): 44–60. 10 mei 2016. <http://www.emeraldinsight.com/doi/pdfplus /10.1108/EUM0000000007075>.

Vakkari, Pertti. “Relevance and contributing information types of search documents in task performance.” *SIGIR* 56.5 (2009): 2-9. 3 juni 2016. <http://dl.acm.org /citation.cfm?id=345512>.

Vakkari, Pertti. “Task Complexity, problem structure and information actions. Integrating studies on information seeking and retrieval.” *Information Processing & Management* 35.6 (1999): 819–837. 5 mei 2016. <http://www.sci encedirect.com/science/article/pii/S030645739900028X>.

Vakkari, Pertti, Nanna Hakala. “Changes in relevance criteria and problem stages in task performance.” *Journal of Documentation*, 56.5 (2000):540–562. 3 mei 2016. <http://www.emeraldinsight.com/doi/pdfplus/10.1108/EUM0000000007127>.

Vakkari, Pertti, Miiko Pennanen en Sami Serola. “Changes of search terms and tactics while writing a research proposal: A longi- tudinal case study.” *IPM* 39 (2003): 445–463. 3 mei 2016. <http://www.sciencedirect.com/science/article/pii/ S0306457302000316>.

Van den Akker, Chiel, et al. “Digital Hermeneutics: Agora and the Online Understanding of Cultural Heritage.” *WebSci '11 Proceedings of the 3rd International Web Science Conference* (2011): 1-7. 7 maart 2016. <http://dl.acm.org/citation.cfm?id=2527039>.

Van den Akker, Chiel, et al. “From Information Delivery to Interpretation Support: Evaluating Cultural Heritage Access on the Web.” *WebSci '13 Proceedings of the 5th Annual ACM Web Science Conference* (2013): 431-440. 7 maart 2016. <http://dl.acm.org/citation.cfm?id=2464491>.

Van den Akker, Chiel. “History as Dialogue: On online Narrativity.” *Low Countries Historical Review* 128.4 (2013): 103 – 117. 1 april 2017. <http://www.bmgn-lchr.nl/index.php/bmgn/article/view/URN:NBN:NL:UI:10-1-110024>.

Van Erp, Marieke, et al. “Automatic heritage metadata enrichment with historic events.” *Museums and the Web* (2011) 10 maart 2016. http://www.museumsandtheweb .com/mw2011/papers/automatic_heritage_metadata_enrichment_with_hi.

Van Passel, Eva, Jasper Rigole. “Fictional Institutions and institutional frictions: creative approaches to open GLAMs.” *Digital Creativity* 25.3 (May 2014): 203-211. 8 mei 2016. https://biblio.ugent.be/publication/6965350>.

Swick, Ivan H. R. “Resource Description Framework (RDF)” *W3C.* 25 februari 2014. RDF Working Group. 3 juni 2016. <https://www.w3.org/RDF/>.

W3C. URI. <https://www.w3.org/wiki/URI>.

Walker, Donald E. *Interactive Bibliographic Search: The User/Computer Interface.* Montvale: AFIPS Press, 1971.

White, Ryen W, Resa. A. Roth. “Exploratory Search: Beyond the Query-Response Paradigm. *Synthesis Lectures on Information Concepts, Retrieval and Services* 1 (2009): 1-98. 25 mei 2016. <http://www.morganclaypool.com/doi/abs/10. 2200/S00174ED1V01Y200901ICR003>.

White, Ryen W, et al. “Supporting exploratory search.” *Communications of the ACM* 49.4 (April 2006): 36-39. 17 mei 2016. <http://portal.acm.org/citation.cfm?doid =1121949.1121978>.

White, Ryen. W, Joemon M. Jose en Ian Ruthven. “Using top-ranking sentences to facilitate effective information access.” *Journal of the American Society for Information Science and Technology* 56.10 (2005): 1113-1125. 27 april 2016. <http://onlinelibrary.wiley.com/doi/10.1002/asi.20203/abstract>.

Wilson, Max L, et al. “From keyword Search to Exploration: Designing future Search Interfaces for the Web.” *Foundations and Trends in Web Science* 2.1 (2010): 1-97. 8 april 2016. <http://dl.acm.org/citation.cfm?id=1734791>.

Wilson, Thomas. D. Models in information behaviour research. *Journal of Documentation*, 55.3 (juni 1999): 249–270. 8 april 2016. <http://www.emeraldinsight.com/ doi/pdfplus/10.1108/EUM0000000007145>.

Yee et al. “Facetted Metadata for Image Search and Browsing.” *SIGCHI’03* (2003)401-408. 10 april 2016. <http://flamenco.berkeley.edu/papers/flamenco-chi03.pdf>.

7. APPENDIX 1

Deze Appendix bevat *user stories* gedaan binnen CLARIAH en DIVE. De user stories zijn tot stand gekomen naar aanleiding van gebruikerstests met de DIVE *tool* met verschillende gebruikersgroepen.

USER STORY 07 "CONSTRUCT & STORE NARRATIVES WHILE NAVIGATING THE DATA"

(DIVE US-end_user-3): For researchers and collection owners DIVE+ with its event-centric approach could provide a great means to build narratives based on collection data.

- Users are able to **refer to narratives**,
    - we need to **save the navigation paths**
    - also entails having a **user account**.

Notes from "old brainstorm" section:

- **Construct narrative:**
    - While browsing the user can select one or more entities that will serve as the **central subject** of the narrative (**Note**: the current collection functionality might be used as a starting point)
    - After selecting entities in this way, the starting point of the narrative is translated into a natural sentence, accompanied by the number of results e.g.:
        - "Tell me the story of *Johan Cruyff* in *Barcelona*" (22)
        - "Tell me the story of *KNIL*" (40)
        - "Tell me the story of *Ruud Lubbers* en *Wim Kok*" (2)

USER STORY 07A: AS A RESEARCHER I WANT TO SAVE MY SEARCH PATH / NARRATIVE SO I CAN LOOK BACK AT THIS FOR FUTURE REFERENCE

This user story originates from the working session in January 2015, where we discussed the creating of a corpus for a researcher.

USER STORY 07B: AS A RESEARCHER I WANT TO SAVE SPECIFIC OBJECTS I FOUND DURING BROWSING SO I CAN BUILD A CORPUS ON MY CHOSEN TOPIC

This user story originates from the working session in January 2015, where we discussed the creating of a corpus for a researcher.

USER STORY 07C: AS A RESEARCHER I WANT TO SHARE MY SEARCH PATH / NARRATIVE WITH OTHER RESEARCHERS SO MY SEARCH PATH CAN BE USED BY OTHERS

This user story originates from the working session in January 2015, where we discussed the creating of a corpus for a researcher.

USER STORY 09 "LINK TO MY NARRATIVE SO I CAN REFER TO IT IN OTHER SYSTEMS"

(DIVE US-end_user-4): For both researchers and collections owners it's important to be able to link to a specific narrative via a URL.

- URLs narratives can be **referred to in publications** such as blog posts or papers or could be stored as a context link in the collection owner's catalogueing system.
- The URL must follow conventions for URL formatting:
    - Readability of URL
    - Style of the URL (Semantic/REST style or regular with URL parameters)
    - Representation of uniqueness of the URL (what kind of ID generator to use)

Moreover to support proper URLs the underlying system must consider the following:

- **Storage of narrative**
    - When a user saves a narrative, the actual data is copied to a narrative storage
- **Storage of narrative queries**
    - When a user saves a narrative, only the user queries are stored
- Support for so called **slugs**
    - User writes his own URL slug
    - System generates a slug from the query parameters

The following are a couple of examples with the aforementioned considerations in mind:

- **Bad example**
    - Not very readable, technical parameters, based on search: http://dive.beeldengeluid.nl?id=2341234zdfdsfsdffasdf&user=harmgot&facets=3,3,12,2&searches=amsterdam+harbour&level=3
- **Better examples**

- Regular style URL with slug generated based on user query and uuid: http://dive.beeldengeluid.nl?narrative=amsterdam_harbour_1930_1950__950e8401-e29b-41d4-a716-446655440030
- REST style URL with just a uuid: http://dive.beeldengeluid.nl/narratives/550e8400-e29b-41d4-a716-446655440000

What the landing pages should look like is described in user story 8.

USER STORY 10 "PUBLISH MY NARRATIVES & SHARE IT WITH OTHERS"

(DIVE US-end_user-8): In order for the sharing of narratives (or the *browsing experience* as it is called on the website help page) to work, it must be possible to refer to each narrative by means of a unique URL (user story 4), which will direct the user to a web page that is:

- **visually attractive**: regardless of user group (stereotypical general public with a low attention span or dry researchers who are content with using a Lynx browser) it seems that most users nowadays would be more likely to spend some time on webpages that are visually attractive/interesting (rather than pages that are not).
- **easy to interpret**: sharing narratives is great, but since narratives in the case of DIVE consist of named entities, event entities, media objects and a handful of semantic relations, interpretations by others who were not present during the deep dive, could be a challenge.

When visualizing the narrative the following could help its interpretation:

- The **starting point** of the narrative, e.g. in the form of "Exploration of events related to the opening of the new Rotterdam train station".
- Any **comments/notes** added by the author (see user story 11).
- **Relation** between the involved entities/events/objects.

USER STORY 08 "PLOT MY NARRATIVE(S) ON A TIMELINE"

(DIVE US-end_user-7): In order to better understand the relation between selected items of (explorative) searches it is important to be able to show these selections on a timeline that can be viewed with no (or minimal) scrolling effort.

- view selected items within a clear time visualization
- consult the metadata (see also user story 6) of each item by e.g. selecting an item.

USER STORY 20 "CREATE COLLECTIONS OF CERTAIN EVENTS AND MEDIA OBJECTS"

(DIVE US-end_user-12): Currently in DIVE it is possible to save search results in a collection.

USER STORY 21 "UNDERSTAND THE SEMANTICS OF PATHS TO IDENTIFY INTERESTING NARRATIVES"

(DIVE US-end_user-9): To aid interpretation it would be useful to be able to inspect the semantics of the relation between the seed object and each found result.

- TODO what relations are currently supported in the DIVE model?

USER STORY 24: AS A RESEARCHER I WANT TO GET SOME BASIC INFORMATION ON MY CHOSEN TOPIC BASED ON THE OBJECTS IN DIVE SO I HAVE A LITTLE MORE CONTEXT

This requirement comes from Annemarie user study, the teams looked at narratives, and also on how to search in DIVE, and they said that if you are exploring a topic you don't know anything about, you should have some data snippets (as Google does) providing basic information about it, for instance, coming from Wikipedia. To create some context about the subject. DIVE as it is now expects you to have some basic knowledge of the objects/topics before you start exploring. If as Chiel said, the idea of DIVE is to support to create questions (not answers) this is ok, but if you want to have some answers you need to have a bit more context. Annemarie says this was tested on students, she is not sure if it applies to researchers as well.

USER STORY 30: AS A RESEARCHER I WANT TO BE ABLE TO SEE OTHER PEOPLE'S NARRATIVE THAT ARE MOST RELEVANT TO MY SEARCH SO I CAN QUICKLY COME TO INTERESTING AND SURPRISING RESULTS

This user story is based on the user study on DIVE with the Data Visualisation class but also mentioned in one of the earlier documents on user stories.
When searching a specific topic for which you have no prior knowledge, it is nice to see the narrative of other people, for they might have found exactly what you are looking for, or at least lead to related interesting objects. This can create more surprising results.

8. APPENDIX 2

OVERZICHT VERSCHILLENDE ONDERZOEKSFASES

In deze Appendix is een tabel te vinden van verschillende fasen van onderzoek geformuleerd door Liliana Melgar. De Tabel wordt Binnen CLARIAH als een referentiepunt gebruikt voor onder anderen het definiëren van informatiebehoeften van onderzoekers uit user studies. Melgar bouwt voort op Bron, van Gorp en de Rijke, Huurdemans en Kamps en De Jong et al. (1535)

Research stages	EXPLORATION	ASSEMBLING[16]	ANALYSIS	PRESENTATION
Researcher's purpose	• Idea generation = What is there?) • Develop a research question or find topic • Perform "preliminary archival research" • Decide if collection is suitable for creating a corpus • Decide sampling strategy	• Purposive information gathering (creating a corpus) • Refine a research question or topic	Same as in Assembling (continuous iterations)	Not supported by CLARIAH tools. However, researcher in this phase "requires integrating original data, the results of analysis, literature, and background material." (Bron et al.)
Information needs	• Muddled topic (diversive curiosity) • Known topic (specific curiosity) • Muddled data (diversive curiosity)	• Known topic • Known item • Known data		
Feelings/Cognition	Exploration: Uncertainty, confusion, doubt. Ends up with Formulation: Clarity	Sense of direction, increased interest		
Information-seeking types	• Exploratory browsing (discovery, recommendations, filtering)	• Focused browsing and directed searching (query (re)formulation, result examination, information extraction)		
Required support	• Providing	• Purposively		

[16] Dit wordt door Bron, van Gorp en de Rijke "contextualization" genoemd (1535)

during browsing/searching	support for online social network and communication (discussing a topic/idea) • Store browsing history & query terms • Store visualizations of search results	select items based on relevance to research question/topic & sampling strategy • Fine-grained filtering		
Required support during visualizing search results	• Observe overall patterns in a collection • Observe links between items in a collection	• Zoom in into specific points of interest		
Required support for annotation	• Bookmarking, commenting & tagging (items + fragments) • Adding floating notes/ideas	• Bookmarking, commenting & tagging (items & fragments) • Create links (&comment) between items or groups of items • Adding notes to items + floating notes/ideas • Connecting to references (background literature)		
Required support from user space + media suite	• Providing support for online social network and communication (discussing a topic/idea) • Store browsing history & query terms • Store visualizations of search results	• Name and store search strategies & browsing paths	• Categorizing tags (creating taxonomies) • Create storylines (proto-narratives) • User visualizations: ways of plotting connections between items (-"by organizing [the material] you create the story") • Connecting to references (background literature) • Shared user spaces (collaborative corpus creation & analysis) • Export items/tags/storylines, etc.	