

Alle termen leiden naar Zandvoort

Het nut van The Internet Archive voor het doen van historisch onderzoek

INTERNET ARCHIVE

A stylized, black silhouette of a classical building facade. It features a triangular pediment supported by four vertical columns. Below the columns is a thick horizontal base. The entire graphic is centered between the words 'INTERNET' and 'ARCHIVE'.

Masterthesis Olaf Verboon 3847810

Master cultuurgeschiedenis Universiteit Utrecht

Docent: Gert-Jan Plets

26 juni 2017

Nederlands Abstract

The Internet Archive is een digitaal archief met een gigantische hoeveelheid bronnen. Begonnen als een verzamelaar van websites, wil The Internet Archive nu alle gedigitaliseerde en *digital born*-bronnen verzamelen. Dit heeft geresulteerd in een collectie van miljoenen archiefstukken en miljarden websites. Een groot deel van de archiefstukken zijn aangedragen door gebruikers. De inhoud van deze archiefstukken is bijna niet onderzocht. Dit onderzoek onderzoekt het door gebruikers aangedragen archief met de vraag of het nuttig is voor historisch onderzoek.

Centraal binnen (digitale) archieven zijn de termen ordening en selectie. Dit wordt de laatste jaren steeds meer gebruikt met het archief een cultureel geheugen te scheppen binnen de samenleving. Aan de andere kant worden persoonlijke digitale archieven steeds belangrijker, welke digitale berichten moeten er van een persoon bewaard worden? Ook hier zijn ordening en selectie van belang, echter door de persoon zelf. In het grote digitale archief is het voor de vindbaarheid van belang dat de archiefstukken ook goed worden beschreven aan de hand van metadata.

Wat vooral opvalt is dat het deze ordening en selectie heel ruim is of ontbreekt in The Internet Archive. Daarnaast zijn de metadata vaak incompleet of foutief. Dit maakt het een ongeschikt archief voor historisch onderzoek. Er is te weinig context en te weinig expertise die bronnen waardeert. Dit kan verbeterd worden door bijvoorbeeld verbindingen aan te gaan met andere professionele digitale archieven. Het is wel een interessant archief om oude websites terug te halen en om eventueel onderzoek doen naar hoe gebruikers het verleden willen herinneren.

English abstract

The Internet Archive is a digital archive with a gigantic number of sources. The archive started as a collector of websites. Later, they wanted to collect all digitalized and digital born sources, like books, audio and video. The collection now includes millions of those sources and billions of websites. A lot of these sources are added to The Internet Archive by users. The content of those sources has almost never been researched properly. This research wants to look at this (user-generated) content and see if the content is useful for historical research.

The two central terms in (digital) archiving are organization and selection. In recent years those two are used in archives to shape a cultural memory in society. Personal digital archives are getting more important as well. Which personal digital messages should be saved? Organization and selection are also important within personal digital archives. They need to be done by the person themselves. In digital archives with a lot of sources findability is also important, this can be improved by applying good descriptions to sources, the so-called metadata.

Organization and selection is mostly absent or too wide in The Internet Archive. The metadata are mostly incomplete or wrong. This makes The Internet Archive not usable for historical research. There is too little context and expertise to evaluate the sources. This could be improved by making connections with other professional digital archives. The Internet Archive is however an interesting research object for old websites and for research into the use of the past by non-professional users.

Inhoud

Hoofdstuk 1 Inleiding.....	5
1.1 Digitale archieven.....	6
1.2 Vraagstelling.....	9
Hoofdstuk 2 De opkomst van het professionele digitale archief	11
2.1 Archiefbeheer in de twintigste eeuw	11
2.2 Postmodernisme in het archief	13
2.3 Digitaal erfgoed.....	16
2.4 Digitale archieven en hun beheerders	18
2.5 Hoofdstukconclusie	24
Hoofdstuk 3 De Gebruiker in het digitale archief	26
3.1 De Filosofie achter de Gebruiker	26
3.2 De gebruiker in het professionele archief	29
3.3 Persoonlijke gebruikersarchieven.....	32
3.4 Digitaal vergeten	34
3.5 Hoofdstukconclusie	36
Hoofdstuk 4 The Internet Archive	37
4.1 De geschiedenis van The Internet Archive en The WaybackMachine	37
4.2 De verzamelmethode van The Internet Archive.....	41
4.3 The Internet Archive in onderzoek en vergelijking	43
4.4 Hoofdstukconclusie	46
Hoofdstuk 5 Nederland in The Internet Archive.....	47
5.1 De zoekmethode	47
5.2 Algemene trends.....	49
5.3 De Persoonlijke Gebruikersarchieven	50
5.4 De actieve gebruikers	53
5.5 Losse archiefstukken	60
5.6 Hoofdstukconclusie	62
Hoofdstuk 6 The Internet Archive voor historisch onderzoek.....	63
6.1 Ordening, selectie en een cultureel geheugen in The Internet Archive	63
6.2 Onderzoek in The Internet Archive.....	66
6.3 Mogelijke oplossingen	68
6.4 Hoofdstukconclusie	70
Hoofdstuk 7 Conclusie	72
7.1 Suggestie verder onderzoek en reflectie.....	74
Literatuurlijst.....	77

Boeken en artikelen	77
Websites en bronnen The Internet Archive	79
Algemene websites en weblinks ter controle van bronnen The Internet Archive.....	79
Bronnen The Internet Archive.....	80
Extra bronnen The Internet Archive met onduidelijke/ontbrekende/gekleurde metadata	81
Gebruikersaccounts The Internet Archive.....	81
Zoektermen The Internet Archive	82
Bijlage 1 tijdslijnen activiteit gebruikersaccounts en zoektermen The Internet Archive.....	83
Tijdslijn 1 Activiteit Gebruikersaccounts en bronnen (4 januari 2003 – 19 juni 2017)	83
Tijdslijn 2 Spreiding zoektermen The Internet Archive (4 januari 2003 – 19 juni 2017)	84

Hoofdstuk 1 Inleiding

Het is een herkenbare situatie. Na het overlijden van een dierbare komen veel persoonlijke foto's en documenten vrij. Foto's en documenten uit een periode van hun leven die jij als nabestaande niet herkent. Ze geven een geschiedenis weer die heel persoonlijk is, maar ook, voor jou, de moeite van het bewaren lijkt. Documenten die zo belangrijk lijken om te bewaren, zelfs voor de volgende generaties of voor een groter publiek dan de directe familie. In de toekomst zullen veel van deze bronnen digitaal zijn. Waar kan je terecht met deze bronnen?

Het plaatselijke archief? Deze heeft wel interesse, maar geen fysieke ruimte voor dit document of heeft al genoeg documenten over hetzelfde onderwerp en de archivaris ziet geen waarde in een extra toevoeging. Waar kunnen deze documenten dan opgeslagen worden? De digitale nalatenschap neem je over op jouw eigen computer, maar daar tref je een grote hoeveelheid documenten aan, waarvan het belang niet helemaal duidelijk is. Daarnaast had de overledene veel pagina's op sociale netwerken als Facebook en Twitter. Echter daar kan je niet bij, want de wachtwoorden daarvan zijn niet bekend.

Gelukkig is er een manier om deze documenten extern op te slaan: het digitale archief. Hier kunnen de gescande of digitale foto's en andere documenten in worden opgeslagen. Je vindt een archief die het in zijn eigen server opslaat, zonder dat daar kosten aan verbonden zijn. Je hebt je eigen gebruikerspagina met jouw eigen collectie van documenten en foto's en de documenten kunnen door middel van links gedeeld worden met wie je wilt. Mochten andere mensen op zoek zijn naar deze bronnen dan kunnen ze die vinden. De naam van het digitale archief waar dit allemaal kan? The Internet Archive.

Deze digitale archieven bieden een nieuwe mogelijkheid voor historisch onderzoek. Welke documenten zijn er te vinden in deze archieven? Hoe goed zijn ze opgeslagen en hoe duidelijk is hun context? Door de mogelijkheid van thuis digitaliseren kunnen nieuwe historische bronnen beschikbaar worden gemaakt, die niet zo snel door een professioneel archief zouden worden opgenomen. De inhoud van deze archieven zijn nauwelijks onderzocht. Dit onderzoek is daartoe een eerste poging.

1.1 Digitale archieven

Digitale archieven zijn het antwoord voor het opslaan van digitale producten. Dit kan van alles zijn: foto's, video's en tekstbestanden. Zowel digitaal erfgoed wat gemaakt is voor commercieel gebruik, denk hierbij naast netwerken ook aan videogames, als persoonlijk digitaal erfgoed. Er zijn veel persoonlijke digitale documenten die niet online staan. Een digitaal archief kan helpen om opslag te bieden voor dit digitale erfgoed en dit doorzoekbaar dan wel inzichtelijk te maken.

Digitaal erfgoed vat meerdere betekenissen samen, wat de literatuur soms enigszins vertroebeld. Soms wordt het vooral bedoeld als bestaand fysiek erfgoed digitaal.¹ Musea en archieven maken hun niet-digitale collectie digitaal. De stukken worden gefotografeerd of ingescand en online beschikbaar gemaakt. Vaak openlijk te bekijken, soms moet er een klein bedrag betaald worden. Dit is anders dan het zogeheten *digital born*-erfgoed wat digitaal is ontstaan. Veel van de huidige professionele archieven bewaren beide soorten digitaal erfgoed.

Er zijn drie kernbegrippen binnen de archieftheorie: ordening, selectie en de maatschappelijke functie. Het digitale archief heeft alle drie grotendeels veranderd. Door de snellere doorzoekbaarheid is de ordening vooral komen te liggen bij de zoektermen die de gebruiker hanteert. De selectie kan door grotere digitale opslagmogelijkheden ruimer worden gemaakt. Deze selectie is veranderd door de scheidslijn gedigitaliseerd erfgoed en *digital born*-erfgoed. Vooral die laatste is een lastige, er wordt digitaal ongelofelijk veel geproduceerd, welke stukken zijn van belang om te bewaren? Hier komt de maatschappelijke functie van het archief ter sprake.

De meest recente postmoderne gedachtegang over het archief is dat het archief in de samenleving moet staan. De verantwoordelijkheid voor selectie ligt niet alleen bij de archivaris, maar ook bij de gebruiker. Het archief moet een breed cultureel geheugen scheppen, voor nu en later en dit kan alleen met hulp van de gebruiker. Hier biedt het digitale archief mogelijkheden. Door zijn online beschikbaarheid is het niet meer nodig om naar het archief zelf toe te gaan. De fysieke vorm van het archief vervalst en de gebruiker kan vanuit elke plek online het archief bezoeken. Dit helpt enorm voor de toegankelijkheid van het archief en de zichtbaarheid van het archief binnen de maatschappij.

¹ Paul Schnabel, 'Voorwoord', in: Henriëke Wubs en Frank Huysmans, *Klik naar het verleden: een onderzoek naar gebruikers van digitaal erfgoed: hun profielen en zoekstrategieën* (Den Haag 2006) 7.

Er is een nieuwe type archiefgebruiker. Digitale archieven worden nu directer doorzocht. De nieuwe gebruiker zoekt door veel archieven naar stukken die hij of zij interessant vindt. Dit betekent dat de stukken vindbaar moeten zijn door gebruikers en de juiste metadata moeten hebben. Metadata zijn de gegevens over de bronnen in het archief, dus datum van vervaardiging, makers, maar ook aanpassingen. Deze metadata werden in fysieke archieven vaak gebruikt door de professionele medewerkers, waardoor een meer gespecialiseerd zoekingsysteem kon worden gebruikt. Nu gebruikers het digitale archief zelf doorzoeken, moet die zoeking overzichtelijk kunnen gebeuren en vraagt om inbreng van de gebruiker zelf.

Bovenstaande is een schets van de problematiek waarmee digitale archieven vanuit professionele instellingen mee te maken hebben. Digitaal erfgoed is in verhouding groter dan alleen het door de professionele instituten aangewezen materiaal. Er zijn grote persoonlijke digitale archieven. Ook deze dienen bewaard en geordend te worden, voornamelijk door diegene die het aanlegt. Welke digitale documenten en berichten vindt hij of zij belangrijk om te bewaren en door te geven aan zijn of haar nageslacht? Er valt te pleiten voor het opstellen van een digitaal testament.²

Het is belangrijk om iemand zelf in staat te stellen zijn of haar persoonlijke digitale archief te beheren. Een sociaal netwerk als Facebook, ondanks zijn historische waarde, kan niet zonder overleg in zijn geheel worden bewaard, dit geldt al helemaal voor meer persoonlijke digitale berichtendiensten zoals WhatsApp en iemands e-mail. Daar kunnen persoonlijke berichten in staan die niet voor iedereen bedoeld zijn. Er is een sterke beweging, ook in de geschiedwetenschappelijke literatuur, die spreekt over het recht om vergeten te worden.³ Het argument is kort samengevat dat samenlevingen in het verleden vaak zijn verbeterd door bepaalde gebeurtenissen en zwarte bladzijdes te vergeten. Internet en digitaal bewaren brengen dit in gevaar. Een beschamende foto kan digitaal namelijk sneller bekeken en gekopieerd worden dan ooit. Privacy en het recht om vergeten te worden moeten in acht worden genomen als iets wordt opgeslagen. Daarom is het nuttiger als een persoon zelf zijn digitale nalatenschap samenstelt.

² S. Pitsillides, J. Jefferies, en M. Conreen, 'Museum of the self and digital death: an emerging curatorial dilemma for digital heritage' in: Elisa Giaccardi (red.), *Heritage and social media. Understanding heritage in a participatory culture* (New York 2012) 56 – 68, aldaar 64-65.

³ Melanie Dulong de Rosnay en Andres Guadamuz 'Memory Hole or Right to Delist? Implications of the Right to be Forgotten for Web Archiving' *Reset Online* 6 (2017), 1- 22, aldaar 1-2.

Dit onderzoek bekijkt hoe digitale archieven die door gebruikers/leken worden samengesteld, verschillen van de digitale archieven die door professionals worden samengesteld. Dit aan de hand van één van de oudste en grootste digitale archieven ter wereld The Internet Archive.⁴ Het archief is in 1996 opgericht met het doel om het internet op te slaan. De oprichter van The Internet Archive zag webpagina's verdwijnen of extreem gewijzigd worden zonder dat dat opgeslagen werd. The Internet Archive ontwikkelde een geautomatiseerde methode om die pagina's in verschillende versies op te slaan en doorzoekbaar te maken doormiddel van de WaybackMachine. Dit is gelukt en er zijn nu al ruim tweehonderdzesentachtig miljard pagina's te doorzoeken.

Naast het opslaan van het internet wil The Internet Archive een nieuwe bibliotheek van Alexandrië te worden. Alles wat door de mens geproduceerd is en digitaal kan worden opgeslagen zou onderdak in The Internet Archive moeten vinden. Hier vallen onder andere boeken, televisieprogramma's, films, foto's en andersoortig beeld- en tekstmateriaal onder. Hier zijn er inmiddels al miljoenen van opgeslagen. Dit met de missie: Alle kennis openbaar voor iedereen. Wat The Internet Archive vooral uniek maakt is dat ze gebruikers hun eigen archieven laten samenstellen. Zover dit onderzoek heeft kunnen vinden is dit het enige digitale archief waar gebruikers direct archiefstukken in kunnen zetten. Hiermee worden de rollen omgedraaid, want veel digitale archieven zijn van professionele instituten die de gebruiker inzetten om hun collectie te promoten. The Internet Archive draait dit gedeeltelijk om.

Ondanks dat The Internet Archive al sinds 1996 actief is, is er weinig inhoudelijk onderzoek naar gedaan. The WaybackMachine wordt vaak gebruikt als tool voor historisch onderzoek: oude webpagina's, die als noot in wetenschappelijke artikelen staan worden via de machine teruggehaald. Verder is er veel geschreven over de manier waarop The Internet Archive technisch werkt, bijvoorbeeld zijn verzamelmethoden. Er is bijna geen onderzoek gedaan naar de inhoud van de andere archieven. Dit is opmerkelijk, want de gebruikersarchieven maken dit een zeer interessant onderzoeksobject. De inhoud van de Nederlandse gebruikersarchieven in The Internet Archive is de focus van dit onderzoek.

⁴ Homepage The Internet Archive <https://archive.org/> (16 juni 2017)

1.2 Vraagstelling

De hoofdvraag van dit onderzoek is ‘Hoe nuttig zijn door gebruikers gevulde digitale archieven in The Internet Archive voor wetenschappelijk historisch onderzoek?’. Dit combineert de hierboven kort aangestipte verhouding tussen de professioneel ingerichte archieven en die door gebruikers ingerichte archieven. Wat doen gebruikers anders dan professionals en hoe heeft dat zijn weerslag op het archief? Het is belangrijk om te vermelden dat The Internet Archive niet als doel heeft om persoonlijke archieven op te slaan, alhoewel het daar wel goede mogelijkheden voor biedt.⁵ De stukken die de gebruikers aanleveren moeten het liefst in een bepaalde vorm gericht zijn op een groter publiek, dus boeken, video’s en historische bronnen die op een grotere belangstelling kunnen rekenen dan persoonlijke documenten. Dat in die opvatting variatie kan zitten, zullen we terugzien in het archief.

Om tot een goede beantwoording te komen van de hoofdvraag worden in hoofdstuk 2 enkele subvragen gesteld over het digitale archief. Als eerste zijn er vragen te stellen over het professionele digitale archief. In welke theoretisch context staat dit archief? Welke termen zijn belangrijk binnen het digitale archief? Wat is de invloed van het postmodernisme op het digitale archief? Wat is de inhoud van het digitale archief? Wat is de rol van de archivaris/professional in het digitale archief? Hoe heeft de digitalisering het archief veranderd? Wat is de status van digitale archieven in Nederland?

Daarna worden er ten tweede in hoofdstuk 3 verschillende subvragen gesteld rondom de gebruiker binnen het professionele digitale archief en persoonlijke digitale gebruikersarchieven. Wat is een gebruiker? Hoe is de rol van de gebruiker veranderd in de digitale wereld? Hoe is de rol van de gebruiker veranderd in het professionele digitale archief? Hoe kan de professional de gebruiker een rol geven in het digitale archief? Is de rol van deze gebruiker zo nuttig? Hoe wordt er in de wetenschappelijke literatuur gedacht over persoonlijke digitale archieven? Wie heeft daar de verantwoording over? Is het handig om digitaal te vergeten en zo ja op welke manier?

Deze vragen over de professionele en persoonlijke digitale archieven geven een opmaat voor de analyse van de archieven in The Internet Archive. Eerst wordt in hoofdstuk 4 kort stil gestaan bij de geschiedenis van The Internet Archive. Hoe is het archief ontstaan en met welke missie? Hoe ziet de algemene inrichting eruit? Wat verzamelen ze en hoe verzamelen

⁵ Aaron Ximm, ‘Active Personal Archiving and The Internet Archive’, in: Donald T. Hawkins (red.), *Personal Archiving: Preserving our Digital Heritage* (Medford 2013) 187 – 214, aldaar 211-212.

ze dat? Wat voor onderzoek is er al gedaan naar de inhoud van The Internet Archive? Welke vorm van digitaal erfgoed wordt er opgeslagen in deze grote archieven?

In hoofdstuk 5 worden de Nederlandse bronnen van The Internet Archive bekeken op historisch nut. Dit wordt gedaan door op verschillende brede historische onderwerpen het archief te doorzoeken. In de eerste plaats de zeer brede term ‘Nederland’, wat ruim 8000 resultaten oplevert met heel diverse bronnen. Daarna is er gekozen op kleinere termen als Tweede Wereldoorlog, Slavernij, Zwarte Piet, Nederlands-Indië en Koloniale oorlog.

Deze termen leveren verschillende resultaten op. Hieruit zijn vooral de verschillende gebruikersaccounts en archieven gefilterd. Welke archieven worden met deze termen gevonden? Hoe actief zijn deze archieven? Er is vooral gefocust op de inhoud van gebruikersarchieven, dit is het meest unieke aspect aan The Internet Archive. Welke invloed hebben zij op de zoektermen? Welke motivaties hebben hun archieven? Hoe zijn hun archiefstukken beschreven? Zijn de archiefstukken juist? Welke kleur geven zij mee aan het archief?

In hoofdstuk 6 worden deze bronnen gekoppeld aan de theoretische discussies uit hoofdstuk 2 en 3, om zo een betoog te houden over het nut van deze gebruikersarchieven voor historisch onderzoek. Voldoen zij aan de eisen die professionals aan digitale archieven stellen? Zijn de bronnen goed genoeg voor gebruik in een historisch onderzoek? Wat kan er verbeterd worden om dit wel zo te maken? Hoofdstuk 7 geeft een uitgebreide conclusie waarbij alle vragen uit de hoofdstukken worden beantwoord en er suggesties worden gedaan voor verder historisch onderzoek, alsmede een reflectie op de eventuele tekortkomingen van dit onderzoek.

Hoofdstuk 2 De opkomst van het professionele digitale archief

In dit hoofdstuk wordt ingegaan op hoe het professionele digitale archief is ontstaan en wat dat betekent voor archieven en archivariissen in het algemeen. Eerst wordt de geschiedenis van archiefselectie en -inrichting in de twintigste eeuw besproken, met als laatste de postmoderne invloed op de archieftheorie. Deze theorie vormt de achtergrond waarop de eerste digitale archieven zijn ontstaan. Daarna worden de eerste digitale archieven beschreven en de problemen waar zij theoretisch en praktisch tegen aan liepen. Om dit goed te begrijpen wordt eerst een korte verhandeling gegeven over de verschillende vormen van de term digitaal erfgoed en de problematiek daaromheen. De focus ligt op Nederlandse digitale archieven, om dit te laten aansluiten bij de casus in hoofdstuk 5. Als afsluiter worden enkele praktijkvoorbeelden gegeven die laten zien hoe anders de digitale archieven zijn ten opzichte van fysieke archieven. Dit als een opmaat voor hoofdstuk 3, waarin we als eerste bespreken hoe gebruikers in professionele digitale archieven een rol spelen.

2.1 Archiefbeheer in de twintigste eeuw

In dit eerste gedeelte wordt vooral de thesis (en later boek) van John Ridener besproken. Hij gaat in op hoe er bij archieven vanaf 1900 wordt nagedacht over het bewaren, ordenen en selecteren van archiefmateriaal. Ridener ziet hier vier periodes in, de chronologische eerste drie worden met behulp van zijn thesis kort toegelicht. De laatste periode heeft betrekking tot de postmoderne invloed, deze worden met andere artikelen uitgebreid besproken, dit is de recentste basis waarin digitale archieven zijn ontstaan.⁶ Er valt voor alle archief-theoretische artikelen te zeggen dat hun basis vooral ligt in ordening en selectie. Hoe dat zo is gekomen bespreekt Ridener in zijn thesis.

De eerste twee periodes die Ridener onderscheidt gaan vooral over ordening. Deze zijn gebaseerd rond twee handleidingen, één van Nederlandse archivariissen uit 1898, de zogenoemde *Dutch Manual* en de andere van Jenkinson uit 1922. Beide luidden een nieuwe periode in. Deze handleidingen zijn vooral geschreven om orde in de archieven te brengen. Veel Europese archieven hadden een groot archief wat alleen geordend was op datum van inname. Er was alleen een chronologische lijst van wanneer een stuk het archief in gekomen

⁶ John Ridener, *From polder to postmodernism: an intellectual history of archival theory* (San Jose 2007) 40 - 42.

was. De *Dutch Manual* was de eerste handleiding die probeerde om archiefstukken te categoriseren, voornamelijk in tijdsperiode.⁷

Jenkinson pleit in zijn handleiding ook erg voor ordening. Waarom hij een nieuwe periode inluidt is omdat hij ook ordent op functie in plaats van tijdsperiodes. Documenten die in tijd van gebruik dezelfde functie hadden of tot een bepaalde overheidsafdeling behoorden, kwamen naast elkaar te liggen in het archief. Naast een tijdsaspect van wanneer het document was vervaardigd, kwam er met Jenkinson een ordening die de nadruk legde op hoe het document was gebruikt. Samen met de *Dutch Manual* legde hij de grondslag voor archieftheorie.⁸

De grootste verandering na deze handleidingen kwam door de Amerikaan T.R. Schellenberg, een archivaris in verscheidene federale archieven in de Verenigde Staten. Schellenberg schreef *Modern Archives: Principles and Techniques* in 1957. Na de Tweede Wereldoorlog was er een toename van (overheids)archieven. Hierdoor kon de oude opvatting van ‘alles bewaren’ wegens fysieke ruimtebeperking niet meer gehandhaafd worden. Schellenbergs *Modern Archives* verschilt ruim van de handleidingen. Met Schellenberg komt volgens Ridener een paradigmaverschuiving, met veel meer focus op de rol van de archivaris en de waardering en selectie van archieven, in plaats van de ordening en beschrijving die in de handleidingen centraal stonden.⁹

Onder Schellenberg verandert vooral de rol van de archivaris. Deze is directer betrokken bij de selectie van archiefstukken. Schellenbergs boek geeft vooral tips voor waardering van deze stukken. Welke zijn belangrijk om te bewaren voor de toekomst en welke niet? De archivaris wordt een selectieprofessional. De archivaris is niet hetzelfde als de historicus. De historicus doet onderzoek naar de bronnen in het archief en beschrijft hun inhoud en context. De archivaris probeert dat al te doen voordat de bronnen het archief in komen. Naast ordening is de selectie het archief binnengekomen.¹⁰

De eerste drie perioden van archieftheorie in de twintigste eeuw focussen voornamelijk op ordening en selectie. Ordening om binnen het archief een duidelijke structuur te hebben, waardoor bronnen makkelijker kunnen worden gevonden en binnen deze ordening relevant tegenover elkaar kunnen zijn. Selectie om de grote hoeveelheid van mogelijke bronnen te

⁷ Ridener, *From polder to postmodernism*, 43-55, 62-63.

⁸ Ibidem, 64-65, 93-94.

⁹ Ibidem, 95 – 97.

¹⁰ Ibidem, 126.

kunnen waarderen. Deze selectie ligt volgens Schellenberg alleen bij de archivaris. De archivaris is niet meer neutraal en drukt zijn stempel op het archief. De kritische reflectie op deze neutraliteit en rol van de archivaris alsmede de rol van het archief in de maatschappij staan centraal bij het postmodernistische denken.

2.2 Postmodernisme in het archief

In zijn laatste periode, de periode van het postmodernisme, onderscheidt Ridener twee belangrijke invloeden. Ten eerste denkt het postmodernisme in paradigma's, dus in periodes waar tijdelijk hetzelfde werd gedacht. Ridener hanteert deze periodes zelf ook, zoals hierboven is beschreven. Deze denkpatronen worden op een goed moment doorbroken. Wat het postmodernisme onderscheidt is dat het deze paradigma's herkent en daar kritisch over kan zijn. De tweede invloed is de digitalisering van het archief, zowel intern binnen het archief als extern naar buiten toe, via digitale archieven op websites. Digitalisering brengt een andere ordening met zich mee.

Één van de eerste besprekingen van het postmodernisme op de archiefwereld zijn twee op elkaar volgende themanummers van het wetenschappelijk tijdschrift *Archival Science* uit 2002, onder redactie van Terence Cook en Joanne Schwartz. De samenvattende inleiding en conclusie die de twee redacteuren hebben geschreven over deze themanummers, zijn erg invloedrijk binnen de recente literatuur, waaronder Ridener, over (digitale) archieven. Deze illustreren het paradigma waarin het archief moet opereren.

In hun introductie brengen ze de veranderende rol van het archief onder de aandacht. Het archief is in de eerste plaats een instituut waar macht een rol speelt. Het is niet simpelweg een pakhuis van oude documenten en bronnen, maar in de eerste plaats een herinneringscentrum. Wat in het archief ligt wordt herinnerd, wat niet in het archief ligt wordt vergeten. De beheerder van het archief moet bij zijn selectie van archiefstukken een keuze maken van wat wordt vergeten. Het grootste gevaar, volgens Cook en Schwartz, is dat als iets eenmaal voor een archief wordt geselecteerd, het een soort vanzelfsprekendheid wordt. De bron is goed genoeg geacht voor het archief en wordt verder niet meer bevraagd op zijn functie. Natuurlijk wordt de inhoud waarschijnlijk wel gebruikt voor onderzoek en aan een kritische vraagstelling onderworpen, alleen niet meer de originele selectie van de bron. Dit is wat Cook

en Schwartz met de themanummers willen bevragen, wat rechtvaardigt de selectie van een bron en hoe kan men als archivaris deze selectie blijven bevragen?¹¹

Hun conclusie geeft antwoord op deze vraag en is opgesteld aan de hand van de artikelen in de themanummers, door andere auteurs. Cook en Schwartz bemerken een verschil tussen theorie en praktijk, niet dat deze totaal uit elkaar liggen, maar een kritische reflectie op beiden is wel noodzakelijk en beide hebben elkaar nodig. Cook en Schwartz zien het beheren van een archief als een performance, waarbij ze Judith Butler volgen. Het archief wordt op een bepaalde manier aan de buitenwereld gepresenteerd, als het ware een toneelstuk, waarbij theorie fungeert als het script van het toneelstuk en de praktijk als de uitvoering van het toneelstuk. De uitvoering kan afwijken van het script en andersom. Dat betekent dat de scripts die bijvoorbeeld Schellenberg en Jenkinson geven over archiveren niet ouderwets zijn, alleen dat er nu een andere uitvoering gevraagd wordt van die scripts.¹²

Wat postmodernisme volgens de twee auteurs heeft veranderd aan het archiefwezen is het recht om bevraagd te worden: op de methode/script en op de uitvoering van dat script. Postmodernisme valt vooral keuzes aan en archivariissen moeten die keuzes kunnen verantwoorden. Ze moeten zich bewust zijn van de scheppende functie die ze hebben in het maken van culturele herinneringen. Om kritische vragen over het maken van die keuzes te kunnen verantwoorden is het belangrijk dat archieven kiezen voor een zo'n breed mogelijke selectie van culturele herinneringen. Niet alleen de bekende mening over een bepaald onderwerp moet ruimte vinden in het archief. Juist de mening die onderbelicht is in het debat zou een plek moeten vinden in het archief, hierdoor wordt de herinnering aan dit onderwerp beter gerepresenteerd in het archief. Archivariissen moeten vooral open zijn over de keuzes die ze maken.¹³

Gabrielle Blais en David Enns voegen de notie toe dat archieven op een bepaalde manier vloeibaar moeten worden. Zij delen de mening dat archieven geen vastgeroeste instituten zijn, maar wijzen op de invloeden van buitenaf op het archief. Niet zozeer de kritische bevraging op selectie en archivariissen waar Cook en Schwartz het over hebben, maar vooral een directere invloed die de buitenwereld heeft op het archief. Sponsors, onderzoekers en gebruikers hebben een grote invloed op hoe het archief bekend staat en wordt gebruikt.

¹¹ Joan M. Schwartz en Terry Cook 'Archives, Records, and Power: The Making of Modern Memory' *Archival Science* 2 (2002) 1 – 19, aldaar 18 – 19.

¹² Terry Cook en Joan M. Schwartz, 'Archives, Records and Power: From (Postmodern) Theory to (Archival) Performance' *Archival Science* 2 (2002) 171 – 185, aldaar 172 – 173.

¹³ Cook en Schwartz 'Archives, Records and Power', 184 – 185.

Archivarissen functioneren niet alleen in een vacuüm, waarbij ze alleen verantwoording hebben naar zichzelf. Daarnaast is het archief niet per definitie het best te begrijpen door de archivaris zelf. De vloeibare staat van een archief betekent vooral dat veel groepen moeten worden aangesproken om dit archief te gebruiken. Archieven moeten zich meer bewust worden van hun publieke functie.¹⁴ In hoofdstuk 3 wordt deze publieksparticipatie uitgebreid besproken.

Bij het nadenken over het archief kan een aantal stappen terug genomen worden, een andere invloedrijke stroming binnen het postmodernisme is nadenken over waar de neiging tot archiveren vandaan komt. Één van de meest invloedrijke ideeën op dit gebied komt van Jacques Derrida. Hij heeft het over *mal d'archive* in het Engels vertaald met *archive fever* of archiefkoorts/archiveringsziekte in het Nederlands. Gebaseerd op het werk van Freud komt volgens Derrida de neiging van de mens tot archiveren uit de neiging van de mens om zichzelf te vernietigen. Volgens Freud heeft de mens de neiging om toe te werken naar de dood, dit ziet Freud vooral in de manier waarop mensen in een psychologische setting de neiging hebben om hun trauma of zware momenten uit hun leven steeds opnieuw te beleven. Dit laat volgens Freud zien dat de mens zijn eigen dood wilt regisseren. De mens heeft de neiging te willen leven, maar ook dood te willen gaan naar zijn eigen instinct.¹⁵

Derrida linkt de door hem genoemde archiefkoorts aan deze doodswil van Freud. Dit kan op meerdere manieren worden opgevat. Doordat de mens werkt naar zijn eigen dood op zijn manier, wil hij zijn leven verzamelen en eventueel zijn archief hem laten overleven. Een zelfgekozen archief stelt de herinnering vast van de overleden persoon. De persoon wordt herinnerd naar zijn eigen instinct.¹⁶ Echter zit het vernietigende aspect ook in het archiveren. Een archief kan in verhouding gelijk worden gesteld aan het leven van een persoon. De persoon heeft controle op zijn eigen archief. Dit betekent dat hij naar lieve lust stukken uit zijn archief kan vernietigen of opnieuw kan beleven. Dit sluit aan bij het in controle zijn over de eigen dood, waar Freud het over heeft.¹⁷ Derrida heeft een erg filosofische benadering van het archief, om Schwartz en Cook in hun redentatie te volgen, heeft Derrida een opzet voor een script van archivering, maar niet eens het complete script. Waarom Derrida praktisch van

¹⁴ Gabrielle Blais en David Enns 'From Paper Archives to People Archives: Public Programming in the Management of Archives.' *Archivaria* 31 (1990) 101-113, aldaar 110.

¹⁵ Jacques Derrida en Eric Prenowitz (vert.), 'Archive fever: A Freudian Impression' *Diacritics* 25 (1995) 2, 9-63, aldaar 14-15.

¹⁶ Marlene Manoff, 'Theories of the Archive from Across the Disciplines', *Libraries and the Academy* 4 (2004) 1, 9-25, aldaar 11-12.

¹⁷ Derrida, 'Archive Fever', 59.

belang is, is vooral omdat hij één van de eerste is die vooral de vraag stelt waarom de mens archiveert?

Een artikel van Tom Nesmith illustreert de rol die digitalisering heeft gespeeld bij het postmoderne archief. Let wel hij schrijft zijn artikel in 2002, vrij vroeg in de digitale revolutie. Nesmith wijst er vooral op dat de computer-gegenereerde bron steeds belangrijker wordt en dat de postmodernistische ideeën zoals hierboven door andere auteurs beschreven van invloed gaan zijn op een nieuwe generatie van archiveren.¹⁸ Waar de *Dutch Manual* en Jenkinson vooral op zoek waren naar een manier om de archieven aan het begin van de twintigste eeuw een andere ordening te geven, is de uitdaging voor archivariissen nu om de digitale stukken een plek te geven. Zowel praktisch in software en onderhoud als representatie. Hoe is na 2002 omgegaan met digitale archieven en hoe vinden de ideeën uit de postmoderne archieftheorie hun weerslag daarin? Dit wordt in het volgende gedeelte beantwoord.

2.3 Digitaal erfgoed

Digitale archieven bewaren in principe digitaal erfgoed. Ze willen toegang bieden tot de stukken in hun archief en deze zijn digitaal te bekijken. Digitaal erfgoed is een paraplueterm die vaak op verschillende manieren gehanteerd wordt. Voordat het ontstaan en gebruik van digitale archieven wordt besproken, is het nuttig om de brede definitie van hun inhoud te bespreken.

Er zijn vaak twee interpretaties van digitaal erfgoed. Paul Schnabel stelt in een voorwoord op een onderzoek naar het gebruik van digitale archieven uit 2006 dat met digitaal erfgoed uitsluitend erfgoed digitaal wordt bedoeld.¹⁹ Erfgoed wat al bestond, maar nu is gedigitaliseerd. Deze definitie van Schnabel is niet juist, maar wel één van de soorten digitaal erfgoed die te onderscheiden is en de soort waar zeker de Nederlandse archieven het meest gebruik van maken. De andere soort is het zogeheten *digital born*-erfgoed, oftewel erfgoed wat alleen in digitale vorm bestaat.

De vorm van digitaal erfgoed kan erg verschillen. Gedigitaliseerd erfgoed beperkt zich voornamelijk tot tekstuele bronnen en fotomateriaal, alsmede het eerst geschoten

¹⁸ Tom Nesmith, 'Seeing Archives: Postmodernism and the Changing Intellectual Place of Archives' *The American Archivist* 65 (2002) 24 – 41, aldaar 41.

¹⁹ Paul Schnabel, 'Voorwoord', in: Henrieke Wubs en Frank Huysmans, *Klik naar het verleden: een onderzoek naar gebruikers van digitaal erfgoed: hun profielen en zoekstrategieën* (Den Haag 2006) 7.

videomateriaal. Dit zijn allen bronnen die in digitale vorm weinig kracht verliezen. Ander groter erfgoed, zoals gebruiksvoorwerpen, industrieel erfgoed of gebouwen met een erfgoedstatus zijn minder makkelijk digitaal te bewaren. Wel na te maken met 3D-modellen en door foto's. Veel auteurs wijzen op het klassieke stuk van Walter Benjamin over kunst en de reproduceerbaarheid daarvan. Erfgoed digitaal, om met Schnabel te spreken, is natuurlijk nooit precies hetzelfde als het erfgoed analoog, het is een digitale reproductie. Een boek is digitaal niet hetzelfde als in fysieke vorm, er hoeven digitaal geen bladzijdes worden omgeslagen en alles wordt van een scherm afgelezen.²⁰

Voor *digital born*-erfgoed geldt dit in mindere mate, de digitale omgeving is de omgeving waar dit erfgoed is ontstaan. Toch is er in de literatuur weinig definitie van wat voor soort erfgoed onder deze *digital born*-categorie valt. Er worden natuurlijk online teksten geschreven, foto's gemaakt en bewerkt, getekend, video's bewerkt, gemaakt en geanimeerd en videogames vallen onder de digitale categorie. Sociale netwerken zoals Facebook, zijn helemaal een klasse apart. Dit is een gigantisch samenraapsel van originele en gedeelde teksten, afbeeldingen en andere digitale producten. Sociale netwerken zijn meer gebruiksproducten met hun eigen vormen. Denk bijvoorbeeld aan Twitter waar berichten maar 140 tekens mogen bevatten. Naast dat er veel tweets worden geproduceerd die historisch interessant kunnen zijn, is het schrijven van een tweet met 140 tekens een uitdaging, wat de kunst van dat schrijven een soort immaterieel erfgoed maakt. Er is genoeg *digital born*-erfgoed, wat nu in digitale professionele archieven maar miniem wordt opgeslagen.

Over gedigitaliseerd erfgoed valt als we Walter Benjamin volgen te zeggen dat het een deel van zijn aura verliest. Zoals boven geschetst een gedigitaliseerd boek is niet hetzelfde als een fysiek boek. Dit geldt in sommige gevallen ook voor digitaal erfgoed. Het beste voorbeeld hiervan zijn videogames. Het in een digitaal archief opslaan van de digitale code van een videogame is niet voldoende om de game te kunnen spelen. Daar is de software voor nodig of de apparaten die dat van origine mogelijk maakte. Ook hier verliest de videogame zijn aura als het niet speelbaar meer is of als het op een andere manier speelbaar is. Als een videogame op een floppydisk staat, kan de floppydisk worden opgeslagen in het archief.²¹

²⁰ Walter Benjamin en Harry Zohn (vert.) 'The Work of Art in the Age of Mechanical Reproduction', in: Hannah Arendt, *Illuminations* (z.p. 1969) 1-26, aldaar 3.

²¹ Joanna Barwick, James Dearnley en Adrienne Muir, 'Playing games with Cultural Heritage: A comparative case study analysis of the current status of digital game preservation', *Games and Culture* 6 (2011) 4, 373 – 390, aldaar 374.

Voor videogames kunnen emulators worden gebruikt. Emulators zijn software die het mogelijk maken om op nieuwere apparaten oudere software te gebruiken, waaronder videogames. Dit is een optie om bijvoorbeeld in een omgeving oude sociale netwerken te laten lopen. Het Amsterdam Museum heeft hier een sterk voorbeeld van, waarbij ze een oud gesloten internetnetwerk uit Amsterdam opnieuw presenteren in hun museum.²² Ook hier is weer het nadeel dat een emulatie een deel van de aura verloren laat gaan, het wordt immers niet gespeeld op het originele platform. Bij het presenteren van *digital born*-erfgoed in een archief moet men nadenken over de ideale en meest originele vorm van dat erfgoed.

2.4 Digitale archieven en hun beheerders

De theoretische basis en theoretische inhoud van digitale archieven zijn besproken, nu rest het ontstaan ervan, hoe is dat zo gekomen? Een antwoord is daarop niet te vinden en ligt grotendeels samen met de manier waarop de rest van de wereld is gedigitaliseerd. Echter vindt de digitalisatie weerslag in de besproken postmoderne ideeën van het archief en het selectieve beleid waar je als archief over moet nadenken. Een digitaal archief past goed bij een postmodern vloeibaar archief waar een kritische bevraging van buitenaf centraal staat. Doordat het archief via een andere weg buiten het fysieke instituut om (via een website) gepresenteerd kan worden, heeft het publiek sneller toegang tot dit archief en heeft veel meer toegang dan voorheen. Daarnaast heeft een digitaal archief een beschermende functie, zeker als de digitale server van het archief niet is ondergebracht in het fysieke instituut zelf. Mocht onverhoopt het fysieke archief verloren gaan, dan zijn er digitale reproducties.

Digitale archieven stellen weer nieuwe vragen aan selectiebeleid. Dit werkt wederom tweeledig voor gedigitaliseerd erfgoed en *digital born*-erfgoed. Digitale opslagruimte is relatief goedkoop en wordt in de toekomst waarschijnlijk steeds goedkoper. De kosten van een digitaal archief zitten vooral in het digitaliseren. Alle niet-digitale stukken moeten worden gescand en gepresenteerd worden in een digitaal archief, het liefst met veel zogeheten metadata eraan. Metadata is de info over het archiefstuk, dus bijvoorbeeld het jaar van vervaardiging van een archiefstuk, vervaardigers, eventuele andere informatie. Daar komt bij dat er in het digitale archief beschreven kan worden wat er met een bron wordt gedaan en hoe hij vindbaar is. Hoe vaak is het bekeken? Hoe vaak gedownload of gedeeld met anderen? Welke trefwoorden zijn er aan verbonden? Ook deze informatie behoort tot de metadata.

²²Homepage Digitale Stad, <http://hart.amsterdammuseum.nl/nl/page/521/re-dds> (20 juni 2017).

Omdat de technologie steeds beter wordt is het zo dat oudere gedigitaliseerde stukken alweer achterlopen in kwaliteit. Daarnaast is digitalisering natuurlijk een door mensen gecontroleerd project, waar in verhouding veel arbeidskosten inzitten, helemaal als scans door verbeterde technologie steeds overnieuw moeten worden gedaan. Gedigitaliseerd erfgoed is in verhouding duurder dan *digital born*-erfgoed om in een digitaal archief te krijgen, want de moeite van het digitaliseren moet worden genomen. Vanwege deze kosten is er een drempel voor selectie van digitalisatie, wat is belangrijk genoeg om in een digitaal archief wederom opgenomen te worden, naast het fysieke archief?

Voor *digital born*-erfgoed geldt dat het digitaliseren niet zoveel kost, maar dat vooral de grote hoeveelheid archiefstukken het archiveren tegenhoudt. Er is een enorme hoeveelheid tekst en afbeelding- en videomateriaal wat zijn oorsprong vindt in de digitale wereld. De grootste uitdaging voor archieven is welk materiaal daarvan belangrijk is. Hier kan tweeledig worden gekozen, omdat het opslaan van dit soort materiaal niet veel kost kan in theorie alles worden opgeslagen, dit betekent niet dat dit moet. The Internet Archive heeft ruim 200 miljard (versies van) webpagina's opgeslagen. Dit is een nauwelijks te omvatten hoeveelheid websites met teksten en foto's. De mogelijkheden zijn er om dit in zijn totaliteit op te slaan. Dit is de grootste uitdaging voor de archivaris nu, uit het nieuwe *digital born*-materiaal een keuze maken, bewaren we praktisch alles of gaan we toch heel erg selecteren?

Hoe zit dit in Nederlandse archieven? Alleen recent is er aandacht geweest binnen Nederlandse archieven voor *digital born*-erfgoed. In 2016 heeft Beeld en Geluid bijvoorbeeld games bemachtigd van één van de eerste Nederlandse game-ontwerpers en is nu bezig met de hierboven beschreven vraagstukken over hoe dat te presenteren.²³ Eigenlijk is het Nederlands digitale archiefbeleid vanaf eind jaren 1990 vooral geweest om bestaande archieven te digitaliseren. Al wordt daar ook al *digital born*-erfgoed genoemd.

Het eerste advies voor het digitaliseren van archieven, met steun van de overheid komt uit 1998: *Alles uit de kast. Op weg naar een nationaal investeringsprogramma digitale infrastructuur cultureel erfgoed*. Dit rapport stelt het nut van digitaliseren van belang en gaat technisch in op hoe de overheid, het ministerie van Onderwijs, Cultuur en Wetenschap (OCW) digitale archieven financieel kan ondersteunen. Het rapport onderscheidt zelfs drie soorten digitaal erfgoed: naast gedigitaliseerd en *digital born* (de laatste wordt in het rapport

²³ 'vroegste Nederlandse videogames in archief van Beeld en Geluid' (versie 12 februari 2016) <http://www.beeldengeluid.nl/nieuws/201602/vroegste-nederlandse-videogames-archief-van-beeld-en-geluid> (19 juni 2017).

gewoon digitaal erfgoed genoemd), ook verslagen over erfgoed digitaal erfgoed. Uit de context valt op te maken dat het hier vooral gaat om foto's van fysiek erfgoed als plekken, gebouwen of voorwerpen.²⁴

Het rapport beschrijft de diversiteit aan archieven die gedigitaliseerd kunnen worden. Niet alleen instituten die nu fysiek de naam archief dragen, zoals het Nationaal Archief, maar ook de collecties van musea en bibliotheken vallen hieronder.²⁵ Het nut van een digitaal archief zit volgens dit rapport vooral in het kunnen delen van de collectie op een openbare manier. Het is vooral een service naar het publiek toe. Digitale archieven maken het mogelijk om publiek sneller te bereiken en het archief beter doorzoekbaar te maken. De grootste waarde is volgens dit rapport dat er sector-overstijgend kan worden gezocht binnen het archief. Digitale zoekmogelijkheden maken het mogelijk om niet alleen binnen één archief te zoeken, maar dwarsverbanden tussen stukken aan te brengen, die er voorheen niet waren. De rest van het rapport gaat in op de procedure waarmee het ministerie van OCW de verschillende archieven financieel wil ondersteunen en wat de archieven zelf zouden moeten bijdragen.²⁶

Van deze financiële steun en advies is goed gebruik gemaakt en inmiddels hebben de meeste grote archieven en instanties in Nederland hun collecties gedigitaliseerd. Soms zijn deze archieven openbaar gemaakt, naar de wens uit het eerste rapport, maar er zijn bepaalde archieven alleen digitaal te onderzoeken binnen het instituut of niet het gehele archief is openbaar online beschikbaar. Nadat deze archieven grotendeels zijn opgezet, reflecteert de literatuur vooral op de veranderingen die deze digitale archieven hebben gebracht voor de instituten en hun medewerkers. Er is een groot corpus aan literatuur over de Nederlandse digitale archieven en hun stand van zaken rond de jaren 2006/2007, dat vooral praktisch en theoretisch kijkt naar hoe het digitale archief functioneert.

Uit 2006 komt het theoretische onderzoek van Annemarie de Jong namens Beeld en Geluid *Het boek van Zand*, zij wil de toekomst van het digitale archief duiden aan de hand van het gebruik van het digitale archief van Beeld en Geluid. Ook wil zij vooral een definitie geven van het digitale archief in verhouding tot het analoge archief.²⁷ De Jong gebruikt *Het boek van Zand* als een metafoor voor het digitale archief. *Het boek van Zand* komt uit een verhaal van de Argentijnse schrijver Borges, een veel aangehaalde literaire schrijver in de

²⁴ W. Adriaans, J van den Berg, L. Breure en A. Melief, *Alles uit de kast. Op weg naar een nationaal investeringsprogramma digitale infrastructuur cultureel erfgoed*. (z.p. 1998) 29-30.

²⁵ Adriaans e.a., *Alles uit de kast* 32.

²⁶ Ibidem, 17-19.

²⁷ Annemieke de Jong, *Het Boek van Zand, Naar een ontologie van het digitale archief* (Hilversum 2006) 3 5.

literatuur rondom digitale archieven, en gaat over iemand die een boek heeft wat constant verandert. Het boek leest nooit hetzelfde. Dit is volgens de Jong een passende metafoor voor het digitale archief, de flexibiliteit.

Wat zijn haar bevindingen? Het archief is niet meer statisch. Het oude archief, volgens de ordening van Jenkinson en de *Dutch Manual*, had zijn stukken individueel beschreven met structuur, context en inhoud, een statische ordening. Dit stond allemaal beschreven op één drager, dit kon variëren van papier tot filmrol. Het digitale document is een stuk flexibeler. Het wordt verspreid over verschillende opslagplekken en kan worden aangepast (denk aan websites). Het belangrijkste om deze aanpassingen te traceren zijn de eerdergenoemde metadata. Naast alle gegevens over het digitale document, moeten de aanpassingen hierin worden weergegeven.²⁸

De flexibiliteit van de stukken vraagt een flexibelere houding van de archivaris. Naast dat deze zich meer moet bezighouden met het up-to-date blijven met de meest recente technologische ontwikkelingen, moet de archivaris actief zijn archief blijven bekijken. Het volstaat volgens De Jong niet meer om een archiefstuk aan te nemen en op te slaan. Doordat de archiefstukken zo flexibel worden gebruikt en gedeeld, moet de archivaris actief blijven kijken wat er met die archiefstukken wordt gedaan.²⁹

De openheid van een archief op het internet is nieuw, geen statig instituut meer waar alles achter dikke muren ligt opgeslagen. Een belangrijke toevoeging is dat een digitaal archief niet meer solo opereert. Het digitale openbare online archief is een archief binnen een web van digitale archieven. De gebruiker heeft een ruime keuze en kan grasduinen en verbanden leggen tussen verschillende collecties. Openbare archiefstukken worden steeds meer vermengd met andere stukken, waardoor er weer nieuwe stukken ontstaan. Zeker in het geval van Beeld en Geluid kan veel van hun audiovisueel materiaal opnieuw worden bewerkt, waardoor er weer een nieuwe video ontstaat. Archieven zouden zich vooral moeten toeleggen op een brede samenwerking tussen hen allen, omdat de gebruiker nu al geen onderscheid meer maakt.³⁰ Sonja de Leeuw trekt specifiek voor audiovisuele archieven in 2012 dezelfde conclusie.³¹

²⁸ De Jong, *Het boek van zand*, 91-92.

²⁹ Ibidem, 92.

³⁰ Ibidem, 93.

³¹ Sonja de Leeuw, 'Het archief als netwerk, perspectieven op de studie van online televisie-erfgoed' *Tijdschrift voor Mediageschiedenis* 14 (2012) 1, 10-28, aldaar 24-26.

De belangrijkste uitdaging voor het toekomstig digitaal archief is volgens De Jong het creëren van een goed cultureel geheugen. Hierin is dezelfde postmoderne vraagstelling te zien als voorheen. De gebruiker wordt hierin steeds belangrijker. Ook de selectievraag, zoals hierboven al aangestipt, is belangrijk. Dat praktisch alles bewaard kan worden, betekent niet dat dat moet gebeuren. In de flexibele omgeving die het digitale archief is geworden moeten vooral de professionals en gebruikers in overleg proberen te komen tot een representatief archief. Selectie is belangrijker dan ooit, volgens De Jong, maar ligt nu niet meer alleen bij de archivaris.³²

Iemand die daar gedeeltelijk anders over denkt is Yola de Lusenet. De Lusenet heeft voor UNESCO in 2002 een charter opgesteld waarin het bewaren van digitaal erfgoed onder de aandacht wordt gebracht. Vooral het opslaan van *digital born*-erfgoed wordt daarin bepleit, volgens eenzelfde model als UNESCO en zijn lidstaten omgaan met materieel en immaterieel erfgoed. Een lijst van UNESCO met internationaal *digital born*-erfgoed (alhoewel veel digitaal materiaal al een internationaal publiek heeft) en organisaties in de lidstaten die het nationale digitale erfgoed promoten. In Nederland is dat het platform Digitaal Erfgoed Nederland (DEN).³³

De Lusenet heeft tevens een interessante bijdrage geleverd aan het vraagstuk over digitale selectie. Anders dan De Jong vindt zij dat het bewaren van alle digitale productie geen directe problemen oplevert. De archivaris moet meer als een tuinman worden, die binnen de grote tuin van het digitale archief de bezoeker wijst op de verschillende mooie bloemen die daar groeien. Echter onkruid mag daar ook groeien. De archivaris moet vooral het onderscheid maken tussen nuttige bronnen, die voldoen aan de vraag van een onderzoeker/gebruiker van het archief en het onkruid voor die specifieke vraag.³⁴

De Jong geeft een goed overzicht van de theoretisch vraagstukken die rondom het (Nederlandse) digitale archief spelen rond 2006. In 2007 komt er namens de overheid een rapport om vooral de archivarissen een beslissende rol te laten spelen bij het maken van de selectie. Dit geldt voornamelijk voor overheidsarchieven, die zeker digitaal, vaak niet heel

³² De Jong, *Het boek van zand*, 94.

³³ Yola de Lusenet, *Preservation of digital heritage, draft discussion paper prepared for UNESCO* (z.p. 2002) 10-11.

³⁴ Yola de Lusenet 'Tending the Garden or Harvesting the Fields: Digital Preservation and the UNESCO Charter on the Preservation of the Digital Heritage', *Library Trends* 56 (2007) 1, 164-182, aldaar 180.

actief werden beheerd. Dit rapport wil daar verandering in zien.³⁵ Een ander probleem wat speelde in de jaren 2006/2007 was van meer praktische aard en dat was het vinden van digitale archieven online door gebruikers. Dit gebeurde in verhouding te weinig, terwijl wel alle moeite was genomen om digitale archieven online aan te bieden. Dit kwam vooral omdat de zichtbaarheid destijds in zoekmachines niet groot genoeg was. Dit lijkt nu, ook met de adviezen tot verdere samenwerking die De Jong en later De Leeuw doen, te zijn verholpen.³⁶

Het meest recent overheidsadvies aangaande digitaal erfgoed komt uit 2015 van het Netwerk Digitaal Erfgoed. Dit werk is getiteld *Nationale Strategie Digitaal Erfgoed* en is vooral gericht op een grotere infrastructuur voor het bewaren van digitaal erfgoed. Deze strategie lijkt het resultaat te zijn van de oproepen die een kleine tien jaar eerder door onder andere De Jong werden gedaan. Het Netwerk Digitaal Erfgoed bestaat uit een aantal grote digitale archieven (waaronder Beeld en Geluid en de Koninklijke Bibliotheek), die samen willen werken aan een overkoepelende infrastructuur. Platform Digitaal Erfgoed Nederland is hier bij betrokken, maar is niet hetzelfde als het Netwerk. De strategie is vooral bedoeld om het nationale digitale erfgoed, zowel *digital born* als de gedigitaliseerde collecties inzichtelijk, bruikbaar en houdbaar te maken.³⁷

Ondanks de problematiek rondom de omvang van de archieven, zijn er genoeg digitale middelen om een archief te doorzoeken. Historisch onderzoek wordt steeds wiskundiger wat dat betreft. Rossaak wijst op het belang van algoritmes in digitale archieven. Met het toepassen van algoritmes kunnen archieven én beter worden doorzocht en het zoekgedrag kan beter in kaart worden gebracht.³⁸ Het toepassen van data-analyse wordt steeds belangrijker in de grote hoeveelheid bronnen die een netwerk van digitale archieven bijeenbrengt.³⁹ Dit toont de flexibele documentomgang die De Jong benoemt. Het is bijvoorbeeld mogelijk, en daar liggen algoritmes aan ten grondslag, om een grote corpus berichten te onderzoeken op

³⁵ K.J.P.F.M. Jeurgens, A.C.V.M. Bongenaar en M.C. Windhorst (red.), *Gewaardeerd verleden: bouwstenen voor een nieuwe waarderingsmethodiek voor archieven* (z.p. 2007) 74-75.

³⁶ Henrieke Wubs en Frank Huysmans, *Klik naar het verleden: een onderzoek naar gebruikers van digitaal erfgoed: hun profielen en zoekstrategieën* (Den Haag 2006) 81-82.

³⁷ Netwerk Digitaal Erfgoed, *Nationale strategie digitaal erfgoed* (z.p. 2015) 19-20.

³⁸ Eivind Rossaak 'Analyzing archives in motion' *Journal of Scandinavian Cinema* 1 (2010) 1, 79 – 81, aldaar 81.

³⁹ Fred Gibbs en Trevor Owens, 'The Hermeneutics of Data and Historical Writing', in: Jack Dougherty en Kristen Nawrotzki (red.), *Writing history in the digital age*, (Ann Arbor 2013) 159 – 170, aldaar 163-164.

emoties. Zeker bij het doorzoeken van digitale archieven van sociale netwerken waar veel persoonlijke berichten staan zijn deze algoritmes erg nuttig.⁴⁰

Een ander belangrijke toevoeging die het digitale archief brengt is het automatisch zoeken. Waar De Lusenet oproept voor een professional die als tuinman uit het archief de ‘mooie bloemen’ selecteert, zo heeft het universiteitsmuseum van Groningen al een bot die dat onder de aandacht brengt. De bot zoekt door het archief en brengt dan via sociale media een archiefstuk onder de aandacht. Om dit goed te doen is het de taak aan de archivaris om dat archiefstuk goed beschreven te hebben.⁴¹ The Internet Archive gebruikt, zoals we later zullen zien, bots om hun collectie te laten groeien. De meeste van de websites, die The Internet Archive bewaart, worden automatisch opgeslagen door dit soort bots.

Het gebruikerspubliek is groter bij een digitaal archief. Een interessant voorbeeld hierin is het Zuid-Afrikahuis. Zij hebben gemerkt, dat sinds zij hun collectie online hebben aangeboden er een veel groter aantal mensen naar hun collectie kijkt. Het huis heeft natuurlijk een internationaal karakter, omdat het huis de relatie tussen Zuid-Afrika en Nederland centraal stelt. Het publiek heeft nu veel sneller toegang tot de collectie van dit kleine museum. Hun gebruikerspubliek is ruim vergroot.⁴² Welke implicaties dat heeft speelt een grote rol in het volgende hoofdstuk.

2.5 Hoofdstukconclusie

Hoe zijn digitale archieven ontstaan en hoe functioneert de professional in deze archieven? Digitale archieven zijn een modern antwoord op de drie basisbegrippen uit de archieftheorie van de afgelopen eeuw: ordening, selectie en verantwoording. Digitale archieven bieden een nieuwe digitale ordening van oud materiaal, de collectie is sneller te bekijken en te doorzoeken. Doordat er geen fysieke ruimte en opslag meer zit tussen de stukken zijn er grote mogelijkheden om stukken via andere onderwerpen te linken dan de originele onderverdeling van stukken in het fysieke archief. Selectie is steeds een belangrijke factor binnen het professionele digitale archief. Wat wordt gedigitaliseerd en hoe worden digitale bronnen het meest ideaal gepresenteerd? Er is nog steeds sprake van een teveel aan bronnen, het digitale

⁴⁰ Erik Boertjes, Almarima Jamakovic en Stephan Raaijmakers, ‘Analyse van sociale media’, *Stator* 3-4 (2011), 44-47, aldaar 45.

⁴¹ Frank den Hollander, ‘Mulerius: erfgoed per tweet: Hoe een robot de digitale collecties doorzoekt...’ *Pictogram* 2 (2016) 4, 6-7, aldaar 7.

⁴² J. Stobbe, C. de Maijer en L. Cavero ‘Van small data naar big audience: Collecties Zuid-Afrikahuis in bewerking (2)’ *Maandblad Zuid-Afrika* 92 (2015) 5, 110-111, aldaar 111.

archief heeft daar zijn valkuilen en behoeft een selectie. Door de openheid van het digitale archief is er steeds meer dialoog aan te gaan over deze selectie. Dit is de meest belangrijke verandering die het digitale archief brengt, het is open, makkelijk bereikbaar en te doorzoeken.

De verantwoording voor selectie is door deze openheid makkelijker te bediscussiëren. Dit vraagt om een openheid van de archivaris over zijn/haar beleid en een meer actievere houding op de selectie. De rol van de archivaris houdt niet meer op na de selectie, de bronnen moeten bekeken blijven worden. Het postmodernisme bracht vooral het inzicht dat het archief een rol speelt in de samenleving, vooral bij het scheppen van een cultureel geheugen. Het digitale archief speelt hier goed op in door zijn openheid en het meest recente idee is dat het archief helemaal geen grenzen meer kent. Aangemoedigd door de gebruikers die digitale archieven inwisselbaar gebruiken is één van de meest recentste voorstellen om vooral de archieven tot één netwerk te maken. Stuk A in archief B past namelijk heel goed bij stuk B in archief A. In Nederland is de meest recente ontwikkeling dat de archieven meer gaan samenwerken, als het ware één archief worden. Hoe de gebruiker daarbij een rol heeft gespeeld en gaat spelen, bespreken we in het volgende hoofdstuk.

Hoofdstuk 3 De Gebruiker in het digitale archief

Na het vorige hoofdstuk waar geschetst is hoe het digitale archief is ontstaan en hoe de professional daarmee omgaat is het nu de tijd om de gebruiker aan het woord te laten. Hoe is de gebruiker veranderd binnen het professionele archief? Hoe legt de digitale mens zijn eigen archief vast en welke valkuilen zijn daarbij? Om te beginnen geven we een kleine filosofie over de gebruiker en hoe hij/zij functioneert op het digitale web en daarmee in digitale professionele archieven, plus enkele kritieken daarop. Daarna wordt ingegaan op de discussies over persoonlijke digitale archief.

De gebruiker speelt een grote rol binnen de casus van The Internet Archive, zij mogen daar zelf archiefstukken aan toevoegen, welke erg persoonlijk kunnen zijn. Dit hoofdstuk bekijkt hoe gebruikers in het algemeen in andere professionele digitale archieven worden toegepast. Daarnaast kijken we naar welke persoonlijke documenten een persoon achter kan laten en welke complicaties daar omheen zitten.

3.1 De Filosofie achter de Gebruiker

Trevor Owens stelt in een klein experimenteel onderzoek op zijn blog de vraag ‘wanneer werden we gebruikers?’ of ‘When did we become users?’.⁴³ Een terechte vraag, in dit onderzoek is de term gebruiker al regelmatig gevallen, zonder daar een definitie aan te verbinden. Owens merkte dat het in de Engelstalige literatuur over de digitale wereld steeds meer ging over *users*, in contrast met bijvoorbeeld vergelijkbare woorden als consument, klant of producent.

Owens stelde, met niet geheel betrouwbare data, een grafiek samen waar het gebruik van het woord ‘user’ in alle boeken (in Google Books) van de twintigste eeuw, werd afgezet tegen het gebruik van de woorden ‘consumer’, ‘customer’ en ‘producer’. Wat in die datagrafiek opvalt is dat vanaf halverwege de jaren 1960 het woord ‘user’ van het minst gebruikte woord explodeert naar veruit het meeste gebruikte woord vijftien jaar later begin jaren 1980. Ook ‘consumer’ en ‘customer’ zijn de laatste jaren gestegen. Dezelfde conclusie kon hij trekken uit

⁴³ Trevor Owens, ‘When did we become users?’ (versie 5 augustus 2011) <http://www.trevorowens.org/2011/08/when-did-we-become-users/> (25 mei 2017).

het doorzoeken van de database van Time Magazine. Owens stelt hier vervolgens twee onbeantwoorde vragen bij: wat is de geschiedenis van de gebruiker en als wij gebruiker zijn van technologie is onze relatie met technologie dan veranderd?

Deze vragen, vooral de tweede, liggen ten grondslag aan hoe gebruikers het digitale archief gebruiken. Er zit een groot verschil tussen gebruiker, producent, consument en klant. Een producent maakt een product vanuit niets en een consument consumeert dat product passief. Een klant vraagt om een product. Natuurlijk zijn een consument en een klant ook een gebruiker van dat product, de producent gebruikt het product ook voor zichzelf of om te verkopen, maar de digitale gebruiker is een geval apart. Bij De Jong in het vorige hoofdstuk is al besproken dat het document in het archief veel minder fysiek is, het kan op veel manieren worden gebruikt en worden gevonden. Het minder fysieke is het onderscheid waar gebruiker zich lostrekt van consument en producent. De gebruiker neemt niet meer als een consument alleen een product tot zich, maar gebruikt dat veel intenser, zet het om in een ander product en is zo een producent.

In de literatuur over de digitale omgeving waarin digitale archieven opereren wordt de invloed van de gebruiker vaak aangeduid als Web 2.0. Web 2.0 geeft natuurlijk aan dat er een Web 1.0 moet zijn. In het algemeen wordt ongeveer tot de eerste opkomst van sociale netwerken rond 2004/2005 de term Web 1.0 gebruikt voor de interactie op het internet en vanaf dan Web 2.0. Web 3.0 is volgens sommige al gestart, al is dat veel minder geaccepteerd.⁴⁴ Het onderscheid tussen 1.0 en 2.0 ligt hem vooral in de gebruikersparticipatie op het internet. De gebruiker van web 1.0 was grotendeels een passieve gebruiker, een consument als het ware. Hij gebruikte het internet om dingen op te zoeken, maar droeg niks bij aan die websites.

Met de opkomst van de eerste internetfora werd Web 2.0 in gestart. Gebruikers begonnen actief bij te dragen aan webpagina's, die niet door hen gecreëerd waren. Ze reageerden op elkaar en begonnen grote digitale netwerken te vormen. Een groot aantal huidige populaire websites zijn hiervan een voorbeeld: Wikipedia als internationale door gebruikers geschreven encyclopedie en Facebook, dat het grootste internationale sociale netwerk wist te worden. Het internet onder Web 2.0 werd een sociale ontmoetingsplaats voor tal van activiteiten, het werd meer van gebruikers, het internet wordt actief in vele digitale vormen gebruikt. Web 3.0 zou weer het samensmelten en overvloeien zijn van die verschillende netwerken. Hierbij valt te

⁴⁴ Bart de Nil en Jeroen Walterus, 'Cultureel erfgoed en web 2.0', in: Bart de Nil en Jeroen Walterus (red.), *Erfgoed 2.0, nieuwe perspectieven voor digitaal erfgoed* (Brussel 2009) 9-17, aldaar 10.

denken aan de recente strategie van de Nederlandse archieven om samen te werken. Er is twijfel of het internet nu al in Web 3.0 opereert, want veel netwerken opereren nog steeds los van elkaar, ondanks dat daar soms dezelfde dingen op worden gedeeld.⁴⁵

De gebruiker van het digitale archief is in de eerste plaats veel actiever dan de oude gebruiker en wil invloed op het digitale archief. Alex Bruns introduceert de term *produsage* om het gedrag van deze gebruikers te duiden. In *produsage* zitten zowel produceren als gebruiken (*Usage*) verweven. De gebruiker wil het internet niet alleen gebruiken, maar ook daaraan toevoegen. In het geval van het digitale archief kunnen (gedigitaliseerde) stukken gebruikt worden in een ander product of een ander soort kunstvorm. Het gebruikersgemak staat voorop, maar ook de mogelijkheid tot hergebruik van digitale bronnen is een grote vraag op het internet.⁴⁶

Daarnaast is de gebruiker een concurrent geworden. Het staat een leek vrij om een digitaal archief aan te leggen of (foutieve) historische informatie te verspreiden via het internet. Dit digitale archief kan zelfs bestaan uit stukken die eerder in een professioneel archief zaten en nu gedigitaliseerd op internet staan. Hetzelfde geldt voor (semi-)professioneel onderzoek doen. Het eerdergenoemde Wikipedia is een lichtend voorbeeld wat dat betreft. Iedereen kan bijdragen aan de kennis van deze encyclopedie, echter wordt die informatie niet altijd op wetenschappelijke wijze toegevoegd. Er wordt niet altijd geciteerd of onderbouwd. Dit betekent niet dat de kennis die op Wikipedia staat enorm foutief is, maar wel altijd extra kritisch moet worden benaderd.⁴⁷

In de recente literatuur over de historicus en de digitale wereld is Wikipedia het meest gebruikte voorbeeld. Hoe zorg je er als historicus voor dat je jouw professionele training kan gebruiken om een online encyclopedie te verbeteren en daarbij niet de ongetrainde personen die willen bijdragen verjaagd? Het meest gegeven advies dat de historicus hiervoor krijgt is vooral de stem van de gebruiker ruimte te geven en te erkennen als kundig. Probeer met de gebruiker een soort van middenweg te vinden tussen enerzijds professionele geschiedschrijving en anderzijds de publieke stem daarin.⁴⁸ In de journalistiek worden

⁴⁵ De Nil en Walterus 'Cultureel erfgoed' 10.

⁴⁶ Axel Bruns 'Produsage: Towards a Broader Framework for User-Led Content Creations' *Proceedings Creativity & Cognition* 6 (2007) 1-7, aldaar 3.

⁴⁷ Robert S. Wolff, 'The Historian's Craft, Popular Memory and Wikipedia', in: Jack Dougherty en Kristen Nawrotzki (red.), *Writing History in the Digital Age* (Ann Arbor 2012) 64-74, aldaar 71.

⁴⁸ Wolff, 'The Historian's Craft' 70.

dezelfde adviezen gegeven over het internet. Ook journalisten zien dat hun gezag veel minder vanzelfsprekend is op een groot netwerk als het internet met veel meningen.⁴⁹

De gebruiker is in de loop der jaren steeds meer een actieve factor geworden bij de productie op het internet. Hij is niet meer alleen een producent en een passieve consument, maar mensen willen in constante verbinding met elkaar staan. In het professionele digitale archief is dit niet anders. Hoe kunnen professionals de gebruiker betrekken bij het beheren en creëren van het archief? Hier ligt weer het scheppen van een cultureel geheugen aan ten grondslag, het gegeven wat vanuit de postmodernistische archieftheorie uit het vorige hoofdstuk naar voren kwam. Een actieve gebruikersbasis lijkt hiervoor een oplossing. Het digitale archief is makkelijk te bereiken en gebruikers kunnen kritisch bijdragen aan de inhoud, doorzoekbaarheid en de presentatie van het archief.

3.2 De gebruiker in het professionele archief

Hierboven is besproken hoe de gebruiker veranderd is in de digitale wereld en wat hij verwacht van het archief, voornamelijk inmenging. Het is belangrijk om hier een klein onderscheid in te maken. Alhoewel de gebruiker veel actiever is geworden in het gebruiken, betekent dat niet dat elke gebruiker dat is. Digitale archieven hebben een toename gezien van veel korte bezoeken. Dit zijn vooral de gebruikers die de flexibiliteit van de vele digitale archieven voor zichzelf gebruiken: één stuk uit dit archief, een ander stuk uit het andere.⁵⁰ Deze gebruikers gebruiken het archief veel minder om te doorzoeken, maar zijn heel specifiek in wat ze daaruit halen en gebruiken.

Het korte gebruik van de digitale archieven is toegenomen, maar dit korte gebruik is niet het meest bruikbare voor de archieven zelf. Hoe gebruikt het digitale archief de gebruiker? De doorzoekbaarheid van het digitale archief is belangrijk om gebruikers aan het archief te binden. Wat opviel in de overgang van het analoge naar het digitale archief is dat professionals en gebruikers andere termen hanteren om het archief te doorzoeken. Professionals hanteren meer vakkundige termen en in de eerste plaats ook de originele fysieke ordening. Gebruikers zijn actief betrokken en bestudeerd om hun termen in het archief toe te

⁴⁹ Shayne Bowman en Chris Willis, *We Media, how audiences are shaping the future of news and information* (z.p. 2003) 7-8.

⁵⁰ Gert Nulens 'Virtuele Participatie', in: Bart de Nil en Jeroen Walterus (red.), *Erfgoed 2.0, nieuwe perspectieven voor digitaal erfgoed* (Brussel 2009) 109 – 120, aldaar 111.

passen, zodat zij kunnen vinden wat zij willen. Hier spreekt wederom de openheid van het archief uit sinds eind twintigste eeuw.⁵¹ Het archief moet er zijn voor de samenleving.

Wat ook is toegenomen is de hoeveelheid metadata rondom archiefstukken. Veel auteurs geven aan de term metadata, vooral de bovengenoemde zoektermen mee die verschillen tussen gebruikers en professionals. De Jong geeft een ruimere betekenis mee. Metadata zijn de data over de archiefstukken, zoals beschrijving, datum en auteur/maker. Het schrijven van deze metadata is het meest intensief voor de archieven en moet goed gebeuren, willen de gebruikers de stukken terug willen vinden. Zeker voor stukken die nog niet gedigitaliseerd waren moest dit opnieuw gebeuren.⁵²

Dit beschrijven van stukken was één van de eerste participerende initiatieven van digitale archieven waar gebruikers (vrijwillig) voor werden benaderd. Dit had een tweeledig doel. Ten eerste werd de grote stapel van gedigitaliseerd erfgoed sneller geïndexeerd. Ten tweede konden de gebruikers al meteen aangeven welke metadata zij prefereerde boven andere. Hiermee kon het verschil tussen aanduiding van professionals en gebruikers worden overbrugd. Dit maakte de gebruikers al meteen een stuk meer geïntegreerd in het digitale archief. Deze zeer actieve gebruikers zijn gedreven, maar niet de grootste groep. Dit is waarschijnlijk de eerdergenoemde specifiek zoekende gebruikers.⁵³

In hoofdstuk 2 werd al kort genoemd dat professionele Nederlandse digitale archieven tien jaar geleden moeite hadden om hun digitale archief zichtbaar te maken op het internet. Hoe kan je je digitale archief promoten zodat meer gebruikers het gaan gebruiken? Het is een suggestie om kunstenaars toestemming te geven om delen van het archief te hergebruiken.⁵⁴ Beeld en Geluid gaf een kunstenaar de opdracht om hun stukken te gebruiken en daarmee iets nieuws te creëren.⁵⁵ Naast dat dit een creatief netwerk aanboort, kunnen deze hedendaagse projecten gebruikt worden om het archief te promoten. Het vergroot de zichtbaarheid van het

⁵¹ Gert Nulens, Eva van Passel en Joke Beyl 'Mag het wat actiever? De gebruiker als prosumer van een multimedia-archief', in: Stoffel Debuysere e.a. (red.), *Bewaring en Ontsluiting van Multimediale Data in Vlaanderen: perspectieven op Audiovisueel erfgoed in het digitale tijdperk* (Leuven 2010) 77 – 89, aldaar 79.

⁵² Bart de Nil 'Eigentijds beschrijven: Het contextualiseren van digitaal audiovisueel erfgoed in archieven' in: Stoffel Debuysere e.a. (red.), *Bewaring en Ontsluiting van Multimediale Data in Vlaanderen: perspectieven op Audiovisueel erfgoed in het digitale tijdperk* (Leuven 2010) 36 – 45, aldaar 38-39.

⁵³ Nulens 'Virtuele Participatie', 111.

⁵⁴ Stoffel Debuysere, 'Onvoltooid Verleden Tijd, Brokstukken als Bouwstenen: Voortbouwen op de ruïnes van het audiovisuele geheugen.', in: Stoffel Debuysere e.a. (red.), *Bewaring en Ontsluiting van Multimediale Data in Vlaanderen: perspectieven op Audiovisueel erfgoed in het digitale tijdperk* (Leuven 2010) 125-139, aldaar 128.

⁵⁵ Kelly Mostert 'Turning Historical Newsreels into Awesome Gif Animations' (versie 4 juli 2016), <https://www.beeldengeluid.nl/en/blogs/research-and-development/201607/turning-historical-newsreels-awesome-gif-animations> (25 mei 2017).

archief in het publieke domein. Een sterk doel voor instituten die het culturele geheugen van de maatschappij willen zijn.

De grote valkuil bij het aanbieden van digitale archieven is de openbaarheid van rechten van de archiefstukken en hoe deze gebruikt kunnen worden. Een goede oplossing hiervoor is het gebruik van Creative Commons. Dit maakt sommige van deze archiefstukken bruikbaar voor hergebruik. Beeld en Geluid is wederom hier een leidend archief in die dit heeft uitgezocht. Zij hebben vooral een voorbeeld gezet met hun openbeelden-project.⁵⁶ Hier worden openbare beelden verzameld (zowel uit hun Nederlandse archief als internationaal) die vrij hergebruikt kunnen worden onder Creative Commons gelijk delen. Het gelijk delen betekent dat de hergebruiker de beelden mag gebruiken en wijzigen als ze maar aangeven dat het originele beeldmateriaal van Beeld en Geluid afkomstig is. Daarnaast is het wel de bedoeling dat het hergebruikte materiaal wederom valt onder dezelfde Creative Commons. Het hergebruikte product mag opnieuw gebruikt worden door anderen, In deze ‘gelijk delen’-vorm mag de hergebruiker ook commercieel verdienen aan zijn remix. Er zijn verschillende Commons die hergebruik of kopiëren onder verschillende (commerciële) rechten vastleggen.

Door dit onderscheid in gebruiksmogelijkheden wordt zowel de rechthebbenden van sommige archiefstukken gerespecteerd als het archief gepromoot. Het professionele digitale archief is niet per definitie openbaar op het internet. Vaak is het volledige digitale archief alleen binnen de muren van het archief te bekijken. Binnen de fysieke muren van de archieven heeft de digitalisatie de doorzoekbaarheid makkelijker gemaakt en kunnen de rechten van de stukken die niet openbaar op het internet kunnen staan worden bekeken in een digitale omgeving. De gebruiker of onderzoeker die naar het archief moet afreizen is niet geheel verdwenen met de introductie van digitale archieven online.

Gebruikers zijn actiever betrokken bij het inrichten van het archief en het (her)gebruiken van dit archief, een vergelijkbaar patroon waar Web 2.0 om bekend staat. Er is ook kritiek op deze aanpak. Andrew Keen is erg tegen de mogelijkheden die Web 2.0 biedt. De steeds grotere amateurgroep die via het internet een podium zoekt voor hun werk is volgens hem niet bevorderlijk voor de kwaliteit. Het sneeuwt het echte talent onder en laat mensen genoeg nemen met producten die maar half zo goed zijn als voor het digitale tijdperk. Keen ziet vooral dankzij Web 2.0 een lagere kwaliteit. In zijn opinie is digitaal erfgoed minder waard

⁵⁶ Het openbeelden-project van Beeld en Geluid, <https://www.openbeelden.nl/> (25 mei 2017).

dan ouder (gedigitaliseerd) erfgoed. De mogelijkheid om oude archiefstukken onder Creative Commons te hergebruiken draagt ook bij aan deze vermindering van kwaliteit.⁵⁷

Een artikel van Terranova sluit aan op deze kritische reflectie van Keen. Zij beschrijft hoe in de Web 2.0-cultuur arbeid steeds minder waard is. Dit is terug te zien in hoe digitale archieven gebruikers gebruiken in hun archief. Het helpen met het identificeren en toevoegen van metadata wordt vaak op vrijwillige basis gedaan door de gebruikers. Dit haalt werk uit handen van mensen die daar in gespecialiseerd zijn.⁵⁸ In het kort valt te zeggen dat de gebruikers vooral een actieve rol hebben binnen het archief en worden aangestuurd door een eveneens actieve professional.

3.3 Persoonlijke gebruikersarchieven

In het digitale domein is de gebruiker steeds meer zelf een archiefmaker. Alle digitale documenten en digitale berichten die iemand produceert blijven vaak opgeslagen. Ook hier gaat weleens wat verloren door kapotte harde schijven en dergelijke, maar veel mensen laten een grote digitale erfenis achter. Dit kan bewust gecreëerd zijn, het aanleggen van een digitaal archief is niet alleen voorbehouden aan professionele instituten. Echter laten veel mensen vooral een computer of ander digitaal medium achter, met een ongeordende hoeveelheid digitale documenten in verscheidene vormen. Het is nuttig voor de latere analyse van de gebruikers en de archieven op The Internet Archive om in te gaan op de dimensies die deze persoonlijke archieven kunnen hebben. De meest centrale vraag is wat bewaar je als persoon niet en wat bewaar je wel?

Wat kan een persoonlijk digitaal archief allemaal omvatten? In de eerste plaats zelf gecreëerde digitale documenten, waarvan de meeste waarschijnlijk niet online staan, maar gewoon op de harde schijf van een computer. Alhoewel er nu een opmars is van de zogeheten Cloud waar documenten automatisch in worden opgeslagen buiten de computer om. De Cloud wordt ook veel gebruikt bij smartphones. Er zijn natuurlijk veel directe afgesloten digitale berichtendiensten, één van de oudste is e-mail. E-mail lay-outs hebben vaak al een archiefachtige indeling, e-mails kunnen in mappen worden gesorteerd. Andere berichtendiensten zijn veel vluchtiger, denk aan chats en WhatsApp-berichten. Dan zijn er

⁵⁷ Andrew Keen, *The Cult of the Amateur: How blogs, MySpace, YouTube and the rest of today's user-generated media are killing our culture and economy* (Londen 2007) 38.

⁵⁸ Tiziana Terranova, 'Free Labor: Producing Culture for the Digital Economy' *Social Text* 63 (2000) 18:2, 33-58, aldaar 37.

nog (semi-)openbare berichten op social media als Twitter en Facebook. Daarnaast nog *digital born*-creaties in video- of fotovorm op platforms als YouTube en Instagram.

Een hele brede collectie aan verschillende digitale documenten en hoe zouden deze gearhiveerd moeten worden voor de gebruiker? Richard Banks ontving bijvoorbeeld de computer van zijn recent overleden vader. Door de grote hoeveelheid aan verschillende documenten en documenten die hij als zoon liever niet van zijn vader wilde zien was het een hele opgave om deze computer te doorzoeken en ordenen. Banks pleit voor makkelijkere tools om digitaal te archiveren. Digitaal beheer helpt ook voor nabestaanden om een goed beeld te krijgen van hun vader, zonder dat ze zelf dingen hoeven tegen te komen die het beeld van de overlevende kunnen veranderen.⁵⁹

Pitsilides e.a. zien ook de noodzaak in van een digitaal testament. Zij merken dat veel digitale berichten toch weinig waarde hebben voor diegene die ze produceert. Een summier bericht op een sociaal netwerk in de strekking van: ‘vandaag spruitjes gegeten.’, heeft weinig betekenis zowel voor degene die het later terugleest als degene die het bericht plaatst. Toch zijn er digitale berichten die volgens Pitsilides e.a. wel nuttig zijn om door te geven aan het nageslacht. Zij stellen een digitaal testament voor, wat mensen in staat stelt om zelf te bepalen welke digitale documenten en berichten belangrijk zijn na hun dood en welke niet.⁶⁰

Zowel deze notie van een digitaal testament als het voorstel van Banks doet denken aan de filosofie van Derrida over archieven. Allebei stellen ze voor om een archief te maken dat vooral na iemands dood van nut moet zijn. Hier ziet de destructieve gedachte van Derrida achter, mensen maken een archief omdat ze hun eigen leven en dood willen beheersen. Wat het digitale tijdperk veranderd heeft aan het archiveren is dat het vaak automatisch gebeurt, zij het niet heel geordend. De neiging van digitale apparaten om alles maar op te slaan, helemaal met de meeste recente Cloud-synchronisaties maakt dat er soms te veel digitale documenten opgeslagen worden. Helemaal als een openbaar gedeelde foto of bericht een eigen leven gaat leiden, waarover hieronder meer.

De mogelijkheden om een persoonlijk digitaal archief bij te houden worden steeds groter. Er kan hier net zo’n presentatie achter zitten als bij een professioneel archief. Het samenstellen van een persoonlijk archief hoeft niet per definitie persoonlijk te zijn. Een

⁵⁹ Richard Banks, ‘Our Technology Heritage’ in: Donald T. Hawkins (red.), *Personal Archiving: Preserving our Digital Heritage* (Medford 2013) 215 – 242, aldaar 217 219 220.

⁶⁰ S. Pitsillides, J. Jefferies, en M. Conreen, ‘Museum of the self and digital death: an emerging curatorial dilemma for digital heritage’ in: Elisa Giaccardi (red.), *Heritage and social media. Understanding heritage in a participatory culture* (New York 2012) 56-68, aldaar 64 65.

interessant artikel van Valois-Nadeau illustreert dit. Zij onderzoekt het persoonlijke digitale archief van een gepensioneerd hockeyspeler uit Canada. Dit persoonlijke digitale archief is gepubliceerd door de zoon van die hockeyspeler en heeft volgens Valois-Nadeau een tweeledige betekenis. Ten eerste is het een eerbetoon van een zoon aan zijn vader en heeft het archief een persoonlijke herinneringsfunctie. Ten tweede, vanwege de publieke status van de man als professioneel ijshockeyspeler, is het archief ook een soort erfgoed voor de stad en hockeyclub waar deze man speelde. Valois-Nadeau zet vooral vraagtekens bij de waarde van dit erfgoed en wie uiteindelijk de verantwoordelijkheid daarvoor zou moeten dragen.⁶¹

Persoonlijke archieven hebben een plaats in het grotere publieke culturele geheugen. Dit was al zo bij fysieke archieven, echter digitale archieven maken het makkelijker om vanuit persoonlijke motivatie een monument op te richten, buiten een fysiek of digitaal professioneel archief om.

3.4 Digitaal vergeten

Een digitaal archief biedt een snel-toegankelijk internationaal cultureel geheugen. Veel websites, apps en sociale netwerken slaan alle berichten en gegevens op die hun gebruikers produceren. Zoekmachines binnen en buiten digitale archieven maken het mogelijk om snel en gemakkelijk te zoeken op onderwerpen en personen. Digitaal wordt heel veel onthouden, meer dan ooit in een fysiek archief mogelijk was. Dit heeft alleen een keerzijde. Wat als er iets wordt opgeslagen wat niet opgeslagen had mogen worden? Wat voor gevolgen heeft dat? Naast digitaal archiveren om onthouden te worden bestaat er het recht om digitaal vergeten te worden.

De bekendste theorie hierover komt uit een boek van Viktor Mayer-Schönberger. Hij begint zijn boek met de sterke anekdote over iemand die solliciteert, maar wordt geweigerd op één oude beschamende feestfoto die op een sociaal netwerk staat. De persoon in kwestie had liever gehad dat deze foto niet openbaar was onthouden en tevoorschijn komt wanneer haar naam werd gezocht op het internet.⁶² Mayer-Schönberger wijst erop dat een samenleving naast herinnert ook vergeet. De digitale wereld heeft ervoor gezorgd dat dit vergeten steeds lastiger wordt, omdat computers geen onderscheid maken tussen data. Hier zit volgens

⁶¹ Fannie Valois-Nadeau, 'Rethinking heritagization through the digitization of familial archives' *Reset* 6 (2017) 1-16, aldaar 12.

⁶² Viktor Mayer-Schönberger, *delete, The Virtue of Forgetting in the Digital Age* (Princeton 2009) 11.

Mayer-Schönberger een weerspiegeling van macht in. Doormiddel van het permanent opslaan worden mensen zo gedoceed dat ze het idee hebben dat ze in de gaten worden gehouden.⁶³

Mayer-Schönberger pleit voor het laten terugkeren van het recht om vergeten te worden. Een houdbaarheidsdatum op digitale gegevens lijkt hem de oplossing, een persoon kan zelf aangeven voor welke data deze houdbaarheidsdatum blijft staan.⁶⁴ Dit is deels omgekeerd van wat Pitsilides e.a. bepleiten over het digitale testament. In plaats van een persoonlijke selectie halen uit de digitale berichten van één persoon en de rest verwijderen na zijn leven, moet de persoon volgens Mayer-Schönberger selecteren uit stukken die vooral verwijderd gaan worden.

In navolging van dit boek is er veel geschreven over de (legale) gevolgen van het recht om vergeten te worden. Belangrijk is om te vermelden dat dit recht wettelijk wordt erkend, bijvoorbeeld door de Europese Unie. Iemands privacy moet worden gerespecteerd, ook digitaal.⁶⁵ Dit recht heeft invloed op de inhoud van digitale archieven. Er verdwijnen namelijk bronnen als er digitale data worden verwijderd. Dulong de Rosnay en Gualamaz geven een ietwat genuanceerd beeld op deze zaken. Ze wijzen er in de eerste plaats op dat copyright claims veel meer laten verdwijnen, dan beroepen op het recht om vergeten te worden. Het recht om vergeten te worden moet wel worden gerespecteerd, ook door digitale archieven. Dit betekent niet dat per se het gehele archief verwijderd moet worden, maar dat er toch een selectie gemaakt van wat (deels) openbaar mag worden gepubliceerd en wat niet. Gelukkig zijn er volgens de twee auteurs al genoeg technologische mogelijkheden om dit snel te doorzoeken.⁶⁶

Het recht om vergeten te worden balanceert wel op een aparte juridische scheidslijn. Een recent artikel uit *NRC Handelsblad* gaat hier uitgebreider op in. Zoals gezegd valt het recht om vergeten te worden onder de privacywetgeving en er zijn een aantal mensen die zich daar op beroepen. Sinds 2014 kan er, door bepaling van de Europese Unie, beroep worden gedaan op Google om zoekresultaten te verwijderen. Tot mei 2017 kwamen er ongeveer zeventhonderdduizend verzoeken binnen. Van de Nederlandse verzoeken werd ongeveer 46 procent gehonoreerd. Er is twijfel, veel verzoeken komen van veroordeelde criminelen en hoe wenselijk is het als hun daad wordt vergeten of via Google onvindbaar wordt gemaakt?

⁶³ Mayer-Schönberger, *delete* 73.

⁶⁴ Ibidem, 113 114.

⁶⁵ Liza van Lonkhuizen, ‘‘Vergeetrecht’ wordt op het web lang niet altijd toegekend’, *NRC Handelsblad*, 11 mei 2017.

⁶⁶ Melanie Dulong de Rosnay en Andres Guadamuz, ‘Memory Hole or Right to Delist? Implications of the Right to be Forgotten for Web Archiving’ *Reset* 6 (2017) 1-22, aldaar 15.

Criminelen hebben volgens het artikel op dit moment een grotere kans om vergeten of onvindbaar te worden gemaakt, dan bijvoorbeeld publieke figuren of mensen met een meer persoonlijke motivatie als de beschamende foto. Dit komt voornamelijk door twee uitspraken van de rechtbank die veroordeelden in het gelijk stelden tegenover Google. Een politicus in publieke functie die in opspraak komt wordt eigenlijk nooit uit de zoekresultaten gehaald, dit vanwege zijn of haar publieke functie.⁶⁷ Er zijn kanttekeningen te plaatsen bij de uitvoering van het recht om vergeten te worden, toch is het zo dat digitale archieven naast aangevuld, ook kunnen worden leeggehaald door gebruikers.

3.5 Hoofdstukconclusie

De gebruiker is centraal komen te staan binnen professionele digitale archieven en binnen de digitale wereld zelf. De gebruiker is veel actiever geworden en produceert en bekritiseert binnen de digitale Web 2.0-wereld. De gebruiker is vluchtig en wil zo snel mogelijk in zijn eigen termen archieven kunnen doorzoeken. Dit biedt nieuwe uitdagingen voor digitale archieven op beheersmatig en technologisch vlak, hoe kan de gebruiker zo snel mogelijk de voor hem/haar nuttigste bron vinden? De gebruiker kan op zijn beurt weer de zichtbaarheid van het archief vergroten, door de archiefstukken creatief te hergebruiken en publiceren of andersoortig te promoten. De grote hoeveelheid gebruikers maakt soms wel dat de bereikbaarheid en kwaliteit van deze promotie lager wordt in een grote digitale wereld met veel stemmen en meningen.

Persoonlijk moeten gebruikers gaan nadenken over hoe zij hun digitale erfenis willen bewaren. Een actieve keuze van digitale stukken lijkt hierbij de beste oplossing. Dit confronteert nabestaanden niet met een grote berg onnodige digitale documenten of documenten die niet voor andere ogen bedoeld waren. Het is makkelijker om sneller een digitaal archief op te zetten, die een herinneringsfunctie kan hebben voor zowel nabestaanden, als bijvoorbeeld de regio waarin deze persoon leefde. Het recht om vergeten te worden is een recente en belangrijke vraag binnen de digitale wereld en de jurisdictie daaromheen. Hoe ga je als archief om met dit recht en vergelijkbare claims als copyright? Een digitaal archief zal toch moeten nadenken over hoe openbaar hun stukken zijn en de aandragers van dat archief moeten respecteren in hun wensen daarover. Hoe dat in The Internet Archive precies is geregeld wordt in het volgende hoofdstuk besproken.

⁶⁷ Van Lonkhuizen, 'Vergeetrecht'.

Hoofdstuk 4 The Internet Archive

Dit hoofdstuk geeft de geschiedenis en algemene inhoud weer van The Internet Archive en zijn functies. Wat staat er wel in dit archief en wat staat er niet in? Hoe werkt The Internet Archive precies? Welk onderzoek is er al gedaan naar The Internet Archive? Wat is het verschil tussen The Internet Archive en bijvoorbeeld andere audiovisuele databanken als YouTube? Dit hoofdstuk vormt de opmaat naar de analyse van het Nederlandse gedeelte van The Internet Archive.

4.1 De geschiedenis van The Internet Archive en The WaybackMachine

Hoe is dit archief ontstaan? Dit komt grotendeels door aan de originele oprichter Brewster Kahle. Hij is één van de eerste internetmiljonairs. Hij maakte een zoekmachine die in 1999 werd verkocht aan Amazon voor een groot bedrag. Kahle heeft zich altijd zorgen gemaakt over het verdwijnen en aanpassen van internetpagina's en dat dreef hem om in 1996 te starten met The Internet Archive en het opslaan van webpagina's. Alhoewel bedoeld als een non-profit, zat er in het begin een klein commercieel doel achter. Om het archief te doorzoeken werd de zoekmachine van Kahle gebruikt. Deze maakte de screenshots van websites en doneerde die vervolgens aan The Internet Archive.⁶⁸ In 2002 waren er op deze manier al ruim 100 miljard webpagina's gearchiveerd, nu in 2017 zijn dat er ruim 286 miljard.⁶⁹

De motivatie van Kahle en The Internet Archive om digitale bronnen te bewaren is vrij onduidelijk. The Internet Archive zelf geeft dat alleen te kennen met hun summiere missie: 'to provide Universal Access to All Knowledge'⁷⁰. Iedereen met internet zou via hun toegang kunnen hebben tot alle kennis in de wereld. Om in de eerste plaats de kennis op het internet te bewaren, want naar hun eigen zeggen deed niemand dat in 1996. Hun slogan/missie bestaat uit toegang tot kennis en niet het brengen van kennis. In een interview uit 2001 stelt Kahle dat hij vindt dat de bibliotheekstatus die The Internet Archive heeft helpt om met deze stukken een lange herinnering te creëren.⁷¹

⁶⁸ Roy Rosenzweig, *Clio-wired: The future of the past in the digital age* (New York 2011) 16 – 17.

⁶⁹ De website van The Internet Archive, <https://archive.org/> (25 mei 2017).

⁷⁰ About The Internet Archive, <https://archive.org/about/> (20 juni 2017).

⁷¹ Lizette Wilson, 'One on one with Brewster Kahle, co-founder of the Internet Archive' (versie 21 oktober 2001), <http://www.bizjournals.com/sanfrancisco/stories/2001/10/22/newscolumn9.html> (19 juni 2017).

Eind 2001 werd The WaybackMachine geïntroduceerd op de website van The Internet Archive. Dit is de tool die het mogelijk maakt om alle webpagina's te doorzoeken. Typ een webadres in de machine en alle opgeslagen versies van dit adres zijn te doorzoeken. Hiermee wordt de verandering van webpagina's weergegeven. Sites die al lange tijd bestaan, zelfs sinds de start van The Internet Archive kunnen zo in oude versies worden weergegeven (zie afbeelding 1). Naast het verdwijnen van webpagina's wordt met de shots van The Internet Archive de aanpassingen aan websites zichtbaar.⁷²



Afbeelding 1, opname in the WaybackMachine van The Internet Movie Database van 1 september 2010. Let op de tijdbalk boven in het midden die alle opgeslagen versies van deze pagina laat zien. Ook is het belangrijk om rechtsboven te zien dat er op deze pagina niet is ingelogd. (Screenshot 25 mei 2017 door auteur).

Is elke website te doorzoeken? Nee, Roy Rosenzweig merkt op, al wat jaren geleden, dat vaak alleen de voorkant van een website wordt gevangen en niet alle links die daaronder vallen. In eerste instantie archiveert The Internet Archive alleen vaststaande webadressen. Sommige webadressen zijn flexibel in die zin. Dit geldt zeker voor doorzoekbare databases zoals andere digitale archieven. Een webpagina geeft dan vaak in zijn webadres de zoektermen van de zoeker weer. Dit soort pagina's worden niet opgeslagen. Daarnaast worden de webpagina's in The Internet Archive vooral opgeslagen door bots, het opslaan is

⁷² The WaybackMachine, <https://archive.org/web/> (25 mei 2017).

een geautomatiseerd proces. Sommige websites blokkeren het gebruik van automatische bots op hun site en worden daardoor niet gearchiveerd.⁷³

Daarnaast geeft The WaybackMachine alleen de voorpagina van een bepaalde website. Een groot netwerk als Facebook bestaat uit duizenden (persoonlijke) pagina's, maar daar is in de WaybackMachine alleen de voorpagina of inlogpagina van te zien door de jaren heen. Het eigenlijke netwerk en de meeste van zijn functies zijn niet te zien (zie wederom afbeelding 1).⁷⁴ The Internet Archive verliest websitedata op het in hoofdstuk 3 besproken recht om vergeten te worden. Mocht iemand zich beroepen op dit recht, dan moet The Internet Archive bepaalde websites gaan verwijderen, ook of juist de oudere versies. Hiervoor gelden dezelfde regels als voor een zoekmachine als Google, zoals beschreven in het vorige hoofdstuk.

Nu een groot deel van het internet is opgeslagen en daar een automatisch systeem achter zit is dan het doel van The Internet Archive bereikt? Nee, ze hebben hun focus breder gemaakt om zoveel mogelijk digitale bronnen samen te brengen. Ze willen nu met hun website de nieuwe bibliotheek van Alexandrië worden. Zoveel mogelijk van alle digitale productie opslaan. Een voorbeeld hiervan is hun 'Livemuziek'-archief, wat opnames en fragmenten van liveoptredens omvat.⁷⁵ Dit zijn niet de grote artiesten der aarde, maar de kleinschaligere optredens van minder bekende artiesten in kleine locaties over de hele wereld.

Hetzelfde willen ze doen voor video, andersoortige audio en boeken. Ook hier ligt de focus op de kleinere producenten. De grote televisie-uitzendingen en dergelijke worden allemaal beschermd door copyright en kunnen alleen op The Internet Archive staan als dit is verlopen. Veel van hun stukken hebben een kleine restrictie voor hergebruik. Sommige stukken kunnen alleen worden geleend als ware het een bibliotheekboek. Niet alles kan zomaar van het archief worden gehaald om daar commercieel in een andere vorm weer geld mee te verdienen. Er wordt duidelijk onderscheid gemaakt tussen de rechten (zie afbeelding 2).⁷⁶ Toch is dit verzamelen goed gelukt en bevatten deze audio-, foto-, video- en tekstcollecties nu miljoenen bronnen, verspreid over ruim 259 duizend collecties.⁷⁷

⁷³ Rosenzweig, *Clio-wired* 17-18.

⁷⁴ Het was eerst mogelijk om de inlogpagina's van Facebook door de jaren heen te bekijken via The WaybackMachine, echter staat Facebook geen robots meer toe (25 mei 2017), dus is er gekozen voor een afbeelding van The Internet Movie Database. Hier is rechtsboven een login te zien die niet kan worden gebruikt, je kan niet The Internet Movie Database als ingelogd account bekijken.

⁷⁵ Het Live Music Archive, <https://archive.org/details/etree> (25 mei 2017).

⁷⁶ 'The State of the Commons', <https://stateof.creativecommons.org/> (25 mei 2017).

⁷⁷ Homepage The Internet Archive, <https://archive.org/> (25 mei 2017).

De lay-out van deze bronpagina's zijn over het algemeen hetzelfde. Bovenaan wordt het document weergegeven. Afhankelijk van de vorm kan dit verschillen. Boeken zijn leesbaar gemaakt en zijn in een boekvorm gegoten. Bladzijdes moeten doormiddel van een muisklik worden omgeslagen. Voor audio en video komen er audio- en videospelers tevoorschijn. Soms kan een bron alleen gedownload worden of moet er doorgedownload worden om deze te bekijken. Onder de bron staan de metadata en beschrijving, algemene data aangeleverd door The Internet Archive (tijd van uploaden e.d.) en extra toevoegingen door de uploader zelf (zie afbeelding 2). Rechts van deze beschrijving staan onder elkaar opties voor downloaden, de collecties waar deze bron zich in bevindt en het gebruikersaccount wat de bron heeft geüpload. Onderaan de pagina staan enkele suggesties voor andere interessante bronnen, gebaseerd op de metadata.



electricsheep-flock-244-32500-8

by Scott Draves and the Electric Sheep

Usage Attribution-Noncommercial 3.0 United States

Topics [electricsheep](#), [alife](#), [art](#), [draves](#), [spotworks](#), [evolution](#), [algorithm](#)

Archive of flock 244 of the Electric Sheep, see <http://electricsheep.org> and <http://scottdraves.com>

Afbeelding 2, zie bij Usage de Creative Commons symbolen. Onderste regel is beschrijving. (Screenshot 25 mei 2017 door auteur)

4.2 De verzamelmethode van The Internet Archive

Hoe verzamelt The Internet Archive zijn materiaal? Zoals hierboven aangegeven wordt een groot deel van de websites verzameld doormiddel van automatische processen. Deze processen worden bots genoemd. Denk hierbij terug aan het voorbeeld van het universiteitsmuseum in Groningen, waar een bot het digitale archief doorzoekt en op gezette tijden een voorbeeld uit het archief twittert. Waar de Groningse bot een stuk uit het archief pakt en promoot via sociale media, doen de bots bij The Internet Archive dit andersom en pakken een website (waaronder Twitter) en stoppen dat in hun archief.

Voor de websites werkt dit goed, alhoewel hierboven al wat bezwaren en valkuilen daarvan zijn besproken. Voor de andere bronnen in het archief wordt dit deels gebruikt. Bepaalde video-livestreams worden bijvoorbeeld door bots gedeeltelijk overgenomen. Dit is niet altijd even nuttig. Om een klein stukje vooruit te lopen op het Nederlandse gedeelte van The Internet Archive, er is bijvoorbeeld een bot die met regelmaat een stuk van RTV Drenthe in The Internet Archive stopt, waarschijnlijk van de livestream op de website van RTV Drenthe.⁷⁸ De upload van RTV Drenthe is altijd ongeveer een uur willekeurig stuk uitzending. Er wordt halverwege ergens ingevallen en verscheidene programma's getoond binnen dat uur, zonder die programma's te benoemen. De titel van het archiefstuk is alleen de tijd en datum van handeling en de boodschap dat het RTV Drenthe betreft. Het is een momentopname die moeilijk te plaatsen is. Nuttiger zou het zijn geweest als de bot een bepaald programma kon onderscheiden, nu is het vrij onduidelijk.

Naast gebruik van bots vraagt The Internet Archive vooral aan gebruikers om bij te dragen aan het archief. Dit kan op verschillende manieren. De meest vanzelfsprekende is dat gebruikers zelf stukken aanleveren in digitale vorm. Hiervoor kunnen ze een account maken en daar dingen op uploaden. The Internet Archive zelf kan assisteren bij het snel uploaden van grote verzamelingen en dergelijke, maar ook bij het scannen en door-leesbaar maken van boeken. De gebruiker zijn materiaal moet vrij beschikbaar zijn en vrij van rechten. De gebruiker heeft een profiel waarop alles wat hij/zij uploadt wordt weergegeven. Dit kan een doel hebben, zoals later in hoofdstuk 5 is te zien of kan bijna net zo willekeurig overkomen als de bots.

⁷⁸ The Internet Archive (hierna TIA), RTV Drenthe 2016-11-04 14:49, https://archive.org/details/archiveteam_videobot_RTV_Drenthe_20161104_1449 (25 mei 2017), merk op dat in de link zelf al staat dat het een videobot betreft.

The Internet Archive is geen solo-platform. Het leent veel van de digitale collecties van anderen. Google heeft bijvoorbeeld het project Google Books, met hetzelfde ideaal van een complete bibliotheek. Dit project focust op het digitaliseren en beschikbaar maken van alle boeken ter wereld. De boeken waar geen copyright meer op zit (voornamelijk uit de negentiende eeuw), staan op The Internet Archive. Deze komen uit de Google Book collectie, maar staan in The Internet Archive.⁷⁹ Dit geldt bijvoorbeeld voor de open beelden van Beeld en Geluid die in het vorige hoofdstuk zijn besproken. Ook deze zijn in 2011 door een archivaris van The Internet Archive toegevoegd. De meest recente daarvan niet.⁸⁰ Hetzelfde geldt voor enkele documenten van de Koninklijke Nederlandse Bibliotheek, die in 2014 veel openbare documenten in The Internet Archive heeft gezet.⁸¹

Zijn er stukken die The Internet Archive ondanks hun nu al gigantische omvang niet verzameld? Ja, in een wat curieus artikel, ooit uitgesproken als presentatie, gaat een medewerker van The Internet Archive, Aaron Ximm, in op hoe de tools van The Internet Archive kunnen helpen bij het beheren van een persoonlijk archief. Persoonlijke archieven zoals die in het vorige hoofdstuk zijn besproken. Ximm ziet in de lay-out en gebruikersvriendelijkheid van The Internet Archive veel nut voor persoonlijke archieven. Als deze software beschikbaar voor het publiek zou zijn, dan zouden veel mensen makkelijk een goed uitziend persoonlijk digitaal archief kunnen maken.⁸²

Echter aan het eind van zijn stuk benoemt Ximm specifiek dat The Internet Archive ondanks hun nuttige tools, niet bedoeld is voor deze persoonlijke archieven. Zij willen focussen op stukken die getoond zijn binnen het publieke domein. Dit kan op kleine schaal, maar een boek of film moet het liefst vertoond zijn aan een publiek, voordat het in The Internet Archive zou kunnen komen. Ximm ziet het liefst dat een ander bedrijf/stichting de tools van The Internet Archive zou kunnen implementeren in software voor persoonlijk gebruik.⁸³ Er valt te twijfelen aan de houdbaarheid van dit onderscheid in stukken, zeker in een massaal archief als The Internet Archive. Daarnaast zijn er enkele voorbeelden te vinden

⁷⁹ Een voorbeeld van een door Google Books gedigitaliseerd boek (NB hier wordt geen archivaris account aangedragen): TIA, De Brusselsche anti-slavernij conferentie 1889-1890, <https://archive.org/details/debrusselschean00vissgoog> (25 mei 2017).

⁸⁰ Een voorbeeld van een Beeld en Geluidarchiefstuk op The Internet Archive: TIA, Internationaal torenspringen, <https://archive.org/details/oi2053> (25 mei 2017), ditzelfde stuk op de openbeelden-website van Beeld en Geluid: Internationaal Torenpringen, https://www.openbeelden.nl/media/2053/Internationaal_torenspringen (25 mei 2017).

⁸¹ Een voorbeeld van een KNB stuk: TIA, Wolff [gegraveerd door Floris Baltasarz] <https://archive.org/details/ned-kbn-all-00014740-001> (25 mei 2017).

⁸² Aaron Ximm, 'Active Personal Archiving and the Internet Archive' Heritage', in: Donald T. Hawkins (red.), *Personal Archiving: Preserving our Digital Heritage* (Medford 2013) 187 – 214, aldaar 210.

⁸³ Ximm, 'Active Personal Archiving' 213.

waarbij er in grote mate sprake is van een mix van persoonlijke en het publieke archiefstukken. Deze voorbeelden worden in het volgende hoofdstuk besproken.

4.3 The Internet Archive in onderzoek en vergelijking

Welk onderzoek is er gedaan naar The Internet Archive? Hoe verhoudt The Internet Archive zich tot andere grote digitale databases als YouTube? Deze laatste vraag is interessant omdat veel kwaliteiten die in de literatuur aan YouTube worden toegedicht, bijvoorbeeld voor het doen van onderzoek naar de inhoud, ook toepasbaar zijn op The Internet Archive. Er zitten wat verschillen tussen YouTube en The Internet Archive, hier wordt later op teruggekomen. Als eerste is het beter om te kijken naar welk onderzoek er al gedaan is met The Internet Archive.

Wat vooral opvalt is dat alleen het websitegedeelte en de WaybackMachine worden gebruikt voor historisch onderzoek. Als via Google Scholar op The Internet Archive wordt gezocht dan refereren veel van de gevonden resultaten naar voetnoten. Dit omdat een oude webpagina is teruggehaald en bekeken. The Internet Archive biedt op die manier echt een onderzoekstool naar oude websites. Veel literatuur gaat vooral over de technologische inrichting van het archief⁸⁴ en het nut van het bewaren van websites voor later historisch onderzoek.⁸⁵

Thelwall en Vaughan doen hier enigszins inhoudelijk onderzoek naar, zij het in 2004. Zij vergelijken de resultaten van internationale webadressen en bekijken of er percentuele verschillen zitten tussen landen in hun representatie in The Internet Archive. Is het duidelijk of een aantal nationale websites van één land minder of minder frequent worden opgeslagen dan nationale website van een ander land? Zij zien hier een licht verschil in, sommige nationale websites lijken percentueel beter gedocumenteerd te worden, echter vinden zij dit niet zorgelijk.⁸⁶

AlNoamany e.a. kijken in een artikel uit 2013 vooral naar hoe mensen de WaybackMachine gebruiken en merken op dat er vooral veel gezocht op nog bestaande

⁸⁴ Elliot Jaffe en Scott Kirkpatrick, 'The Architecture of the Internet Archive' *SYSTOR '09 Proceedings of SYSTOR 2009: The Israeli Experimental Systems Conference* artikel 11 (2009) 1-10.

⁸⁵ Eli Edwards, 'Ephemeral to Enduring: The Internet Archive and Its Role in Preserving Digital Media' *Information Technology and Libraries* (2004) 3-8.

⁸⁶ Mike Thelwall en Liwen Vaughan, 'A fair history of the Web? Examining country balance in the Internet Archive' *Library & Information Science Research* 26 (2004) 162 – 176, aldaar 173 174.

websites. Dit vooral om dode links te doen herleven.⁸⁷ Thelwall en Vaughan concluderen hetzelfde voor academisch publiek. De links in voetnoten van oude artikelen worden voornamelijk teruggehaald, mochten deze niet meer werken.⁸⁸ Er is echt nauwelijks onderzoek te vinden wat echt ingaat op de inhoud van The Internet Archive, zeker de inhoud buiten de websites om. Dit onderzoek probeert hier een eerste grotere stap mee te zetten.

Er is één inhoudelijk onderzoek gevonden, bijvoorbeeld in de vorm van een thesis. Anya Boyle heeft in 2016 de archiefcollectie in The Internet Archive over 11 september 2001, de aanslag op het World Trade Center in New York, nader bekeken. Ook zij focust vooral op de praktijk van webarchivering. Wat zij concludeert aan de hand van het 11 september-archief is dat het vooral een groot onduidelijk archief is. Het is niet per definitie bedoeld voor onderzoek, maar vooral als een archief voor overlevenden van de aanslag om hun ervaring te delen. Ook is de ordening van het archief niet heel duidelijk.⁸⁹ In hoofdstuk 5 en 6 bekijkt dit onderzoek in hoeverre dit overeenkomt met de Nederlandse archieven.

Er zijn andersoortige commerciële archieven of databases zoals YouTube. Waarop verschilt YouTube vooral van The Internet Archive? Het verschil zit hem op twee punten. Ten eerste YouTube's naamsbekendheid als videoplatform is een stuk groter dan dat van The Internet Archive. Dit blijkt uit de hoeveelheid gebruikers en video's. YouTube heeft bijna een miljard gebruikers, die vooral passief kijken. Die kijken gezamenlijk bijna honderden miljoenen uren video per dag. YouTube zelf geeft niet aan hoeveel video's er op hun site staan, maar geven wel aan dat er al op 400 miljoen video's copyright claims zijn gemaakt.⁹⁰ Vergelijk dit met de 3,3 miljoen en 1,3 miljoen video's die The Internet Archive aangeeft in zijn collecties video en televisie. YouTube is veruit de grootste, al is de totale omvang van video's op YouTube niet bekend.

Deze naamsbekendheid en grootte van YouTube valt samen met het tweede verschil, namelijk de commerciële doelstelling van YouTube. YouTube werkt met adverteerders en investeert in de makers van video's. YouTube vergelijkt zichzelf meer met tv-netwerken, dan met een echt archief. Videomakers met veel kijkers krijgen door adverteerders geld voor hun

⁸⁷ Yasmin AlNoamany, Ahmed Alsum, Michele C. Weigle en Michael L. Nelson 'Who and what links to the Internet Archive' *Int J Digit Libr* (2014) 14, 101 – 115, aldaar 113.

⁸⁸ Thelwall en Vaughan, 'A fair history of the web' 173.

⁸⁹ Anya Boyle, *Archiving the Internet, what does it mean in practice?* (Leiden 2016) 34 35.

⁹⁰ YouTube Statistics, <https://www.youtube.com/yt/press/statistics.html> (20 juni 2017).

video's en zijn deels in dienst bij YouTube.⁹¹ The Internet Archive is een non-profit archief. Deze commercialiteit is het grootste verschil tussen de twee platforms.

Toch wordt YouTube als een archief gebruikt om filmpjes uit de privésfeer op te zetten. Wat is de waarde van zo'n videoarchief en voor een videoarchief als The Internet Archive? Ook hier is het spanningsveld tussen professional en gebruiker terug te vinden. Robert Gehl ziet YouTube vooral als een te groot archief waar te weinig selectie op wordt gedaan. Dit in lijn met De Lusenet en De Jong over de professional als tuinman dan wel selecteur van het archief. Gehl prijst YouTube met het verzamelen van zo'n divers archief aan bronnen, maar is erg sceptisch over de rol van de gebruiker daarbij. Door de grote hoeveelheid gebruikers is er geen selectie meer van bronnen. Daarnaast door de commerciële kant van YouTube wordt er onderscheid gemaakt tussen video's die met een commercieel oogmerk zijn gemaakt of worden verwijderd wegens copyright-rechten en video's die als een archiefstuk op YouTube zijn gezet.⁹²

Gehl is vooral sceptisch over YouTube als archief en vindt vooral dat er te weinig curatie plaatsvindt, de gebruiker en de commercie zijn te veel bepalend. Dit geldt natuurlijk voor de 4,6 miljoen video's die op The Internet Archive zijn te vinden en hetzelfde argument kan worden gemaakt voor de andere bronnen. De gebruiker heeft daar voornamelijk invloed op. Commercie heeft een zekere rol, echter alleen als er sprake is van copyrightclaims, er zitten geen adverteerders achter, behalve als de gebruikers dat zelf zijn.

Sheenagh Pietrobruno is een stuk positiever over YouTube als archief. Hij ziet dit videoarchief vooral als een middel om immaterieel erfgoed te promoten. Video's van traditionele handelingen of terugkerende gebeurtenissen vinden een plek en een zekere promotie op YouTube. Dit geldt voor immaterieel erfgoed wat gedeeltelijk controversieel is en niet wordt erkend door UNESCO of als nationaal erfgoed. Video's kunnen dienen als een soort gedigitaliseerd immaterieel erfgoed. In tegenstelling tot Gehl is de gebruiker in Pietrobruno's artikel een stuk positiever belicht. Door zijn neiging tot vastleggen wordt deels onzichtbaar immaterieel erfgoed zichtbaar gemaakt en gepromoot.⁹³ Dit is natuurlijk een goede functie die de collecties van The Internet Archive kunnen hebben. Als het doel van The Internet Archive

⁹¹ YouTube Statistics, <https://www.youtube.com/yt/press/statistics.html> (20 juni 2017).

⁹² Robert Gehl, 'YouTube as archive, who will curate this digital wunderkammer?' *International journal of cultural studies* 12 (2009) 1, 43-60, aldaar 56.

⁹³ Sheenagh Pietrobruno, 'YouTube and the social archiving of intangible heritage' *New Media & Society* 15 (2013) 8, 1259 – 1276, aldaar 1272 – 1273.

is om zoveel mogelijk gedigitaliseerd en *digital born*-erfgoed te verzamelen, dan past vastlegging van immateriële tradities daaronder.

4.4 Hoofdstukconclusie

The Internet Archive is in de eerste plaats opgericht om het internet te bewaren en om toegang te geven tot alle kennis van de wereld. Dit lijkt grotendeels gelukt, er zijn ruim 200 miljard (versies van) webpagina's gearchiveerd en te bekijken via de WaybackMachine. Het volgende doel is om een compleet archief van alle gedigitaliseerde en *digital born*-bronnen te worden. Ook hier zijn al enkele miljoenen stukken in audio, video en tekst beschikbaar. De verzamelmethode om deze bronnen bij elkaar te brengen worden aan gebruikers en bots overgelaten. Bots hebben als nadeel dat zij soms geblokkeerd worden bij dit verzamelen en niet selectief of de volledige bron archiveren, waardoor archiefstukken onduidelijk kunnen worden, denk aan het willekeurige uur RTV Drenthe. Gebruikers, ook binnen YouTube, hebben als voordeel dat ze oog hebben voor tradities waar professionals niet aan zullen denken en als nadeel dat er vooral een heel willekeurige en niet op waarde geschatte collectie ontstaat.

In hoeverre dit geldt voor het Nederlandse gedeelte van The Internet Archive wordt in het volgende hoofdstuk besproken. Inhoudelijk is The Internet Archive nauwelijks aan onderzoek onderworpen. Veel literatuur wijst vooral op het nut van het terugzoeken van oude links via de WaybackMachine, maar gebruikt The Internet Archive vooral als middel en niet als inhoudelijk archief. Het gedeelte buiten de websites om zijn nauwelijks bestudeerd. Boyle doet dat als enige en vindt vooral in het door haar bekeken archief een erg persoonlijk gemotiveerd en slecht gestructureerd archief terug.

Hoofdstuk 5 Nederland in The Internet Archive

In dit hoofdstuk wordt gekeken naar de Nederlandstalige archiefstukken binnen The Internet Archive. Deze stukken zijn bijna allemaal aangeleverd door gebruikers. In dit hoofdstuk worden deze gebruiksarchieven voornamelijk besproken en deels geanalyseerd, in het volgende hoofdstuk worden zij vergeleken met de theorieën en opmerkingen uit de voorgaande hoofdstukken, om zo een conclusie te vinden over het nut van deze stukken en The Internet Archive voor het doen van historisch onderzoek. Als eerste wordt kort aandacht geschonken aan de methode waarmee deze archiefstukken zijn gevonden, daarna worden algemene trends in het totale archief en enkele individuele bronnen en gebruikers uitgelicht die typerend zijn voor de inhoud van het archief.

5.1 De zoekmethode

Dit onderzoek beoogt een overzicht te geven van de Nederlandse inhoud van The Internet Archive en om vraag te stellen hoe nuttig deze inhoud is voor wetenschappelijk onderzoek? In de eerste plaats is binnen het totale archief van The Internet Archive gezocht op het woord ‘Nederland’ en dit is gesorteerd op ‘Date archived’.⁹⁴ Dit laatste om in chronologische volgorde van jongste archiefstuk naar oudste archiefstuk enige trends op te merken. Vooral de activiteit van accounts is op deze manier te zien (zie tijdlijn 1 en 2). Er moet worden gezegd dat The Internet Archive een actief archief is, er komen in een gestaag tempo archiefstukken bij. Bij het eerste doorkijken van het archief medio april 2017 op het onderwerp Nederland kwamen er 8358 archiefstukken uit, dit aantal is tijdens het schrijven van dit onderzoek toegenomen.

Om het historisch onderzoek naar de bronnen in dit archief te simuleren is er gezocht op enkele historische/controversiële brede onderwerpen binnen de Nederlandse geschiedwetenschap en maatschappij. Dit zijn de volgende onderwerpen met tussen haakjes de hoeveelheid stukken die worden gevonden daarbij (stand april 2017): Tweede Wereldoorlog (285), Slavernij (36), Zwarte Piet (33), Nederlands-Indië (24) en Koloniale oorlog (10). Binnen deze onderwerpen wordt bekeken wat voor soort bronnen er het meest

⁹⁴ <https://archive.org/search.php?query=Nederland&sort=-publicdate> (6 juni 2017).

naar voren komen en wie die bronnen heeft aangeleverd. Hoe interessant voor onderzoek en duidelijk zijn deze bronnen?

Dit onderzoek kijkt vooral naar algemene trends en opvattingen binnen deze zoekresultaten. Niet elke bron, zeker niet uit de 8000+ resultaten op de tag ‘Nederland’, is in detail bekeken. Vooral de algehele trends en opvallende titels zijn bekeken en gewaardeerd in het licht van de theorie uit de eerdere hoofdstukken, de koppeling daarvan vindt plaats in hoofdstuk zes.



Tijdslijn 1. Activiteit in hoofdstuk 5 besproken gebruikersaccounts en bronnen sinds eerste upload met zoekterm ‘Nederland’ op 4 januari 2003 tot 19 juni 2017. Opvallend is vooral de verhouding stukken en tijd. Bitterburg heeft het op één na grootste archief maar is in verhouding heel kort actief op de site. Hans Steen is het kortst actief geweest, maar heeft meer stukken tot nu toe dan Joep Zander in negen jaar activiteit. Voor grotere versie zie bijlage 1. Tijdslijn gemaakt door auteur.



Tijdslijn 2. Verhoudingen zoektermen door de tijd heen, van 4 januari 2003 tot 19 juni 2017, resultaten ook gepeild rond 19 juni. Hier valt vooral de spreiding op in tijd. ‘Zwarte Piet’ geeft niet veel resultaten (35), maar heeft wel het vroegste stuk binnen de specifieke zoektermen. Koloniale oorlog heeft één hele korte periode van stukken binnen een jaar, terwijl de rest veel spreiding toont. Er moet bij vermeld worden dat er waarschijnlijk tijdens het schrijven van dit onderzoek metadata zijn aangepast. Slavernij gaf in april 2017 nog 36 resultaten, op 19 juni 2017 zijn dit er 50 terwijl er sinds januari 2017 geen nieuwe stukken zijn toegevoegd. Voor grotere versie, zie bijlage 1. Tijdslijn gemaakt door de auteur.

5.2 Algemene trends

Wat valt er op in het grote overzicht rond de term ‘Nederland’? Ten eerste is niet alles afkomstig uit Nederland of Nederlandstalig. Het nadeel van een algemene zoekterm is dat je alles krijgt wat gelieerd is aan het woord Nederland. Hier valt bijvoorbeeld het dorpje Nederland in Colorado in de Verenigde Staten onder.⁹⁵ Hetzelfde geldt voor anderstalige bronnen die Nederland als woord gebruiken, zo zijn er Spaanstalige en Duitstalige bronnen.⁹⁶ De specifiekere zoekterm ‘Koloniale oorlog’, die hierboven benoemd wordt, valt af als Nederlandstalige zoekterm. De tien resultaten (negen audiobestanden en één boek) die daar bij tijd van schrijven uit komen hebben alle betrekking op Zuid-Afrika en zijn in het Afrikaans opgesteld.⁹⁷

De bronnen binnen de zoekterm ‘Nederland’ zijn erg divers, wat niet vreemd is met zo’n brede zoekterm. Andere meer specifieke zoektermen hebben hier ook last van. Zoals in het vorige hoofdstuk al genoemd zijn er een aantal Nederlandse digitale archieven op The Internet Archive gezet, dit zijn specifiek openbare collecties van de Koninklijke Nederlandse Bibliotheek (KNB) en Beeld en Geluid. Tevens zijn veel van de oudere copyrightvrije boeken uit Google Books op The Internet Archive gezet. Bij de zoekterm ‘Slavernij’ komt men bij de 36 resultaten een aantal negentiende-eeuwse boeken in digitale vorm tegen.⁹⁸ Deze zijn allemaal geschreven rondom de afschaffing van de slavernij in Nederland of vertalingen/luisterboeken van oudere bronnen. Niet alle boeken komen direct van Google Books. Het is belangrijk om te onthouden dat deze boeken van Google Books komen en door hun gedateerdheid geen rechten meer hebben, maar waarschijnlijk niet door Google Books op The Internet Archive zijn gezet.⁹⁹

Wat opvalt aan de uploads van de KNB en Beeld en Geluid is dat deze allemaal in dezelfde periode plaats vinden, veel uploads in een kort aantal dagen. Ook hun archieven zijn niet door henzelf in The Internet Archive gezet. De suggestie wordt gewekt dat er een

⁹⁵ The Internet Archive, (hierna: TIA), Inside Boulder News – Flood Update https://archive.org/details/Inside_Boulder_News_-_Flood_Update (6 juni 2017).

⁹⁶ TIA, 0619 Historische Tatsachen Nr. 53 Udo Walendy Enstellte Geschichte (1992, 44 S. Verboten) <https://archive.org/details/0619HistorischeTatsachenNr.53UdoWalendyEntstellteGeschichte199244S.Verboten> (6 juni 2017) en TIA, Jorge Valdes, Radio Nederland Wereldomroep <https://archive.org/details/RNWMixEdicionEspecialJorgeValdes> (6 juni 2017).

⁹⁷ <https://archive.org/search.php?query=Koloniale%20oorlog> (6 juni 2017).

⁹⁸ <https://archive.org/search.php?query=Slavernij> (6 juni 2017).

⁹⁹ TIA, De slavernij in Suriname; of, Dezelfde gruwelen der Slavernij: die in de “Negerhut” geschetst ... <https://archive.org/details/deslavernijinsu00wolbgoog> (6 juni 2017).

medewerker of automatisch proces van The Internet Archive is aangewezen/aangezet die in korte tijd de openbare collectie van deze instanties op The Internet Archive heeft gezet. Voor Beeld en geluid was dat midden augustus 2011, voor de KNB eind oktober 2014.¹⁰⁰ Beiden omvatten een groot aantal stukken en worden actief geüpload over een aantal dagen. Zeker door de grote omvang binnen de korte tijd, lijkt dit een automatisch proces te zijn. Er is sinds de maanden in 2011 en 2014 door beiden niks meer toegevoegd. Het is bij tijd van schrijven een eenmalige actie.

Automatische processen of bots zijn al uitgebreid in hoofdstuk 4 beschreven als verzamelmethode van The Internet Archive. Hiervan is vooral het account wat steeds stukken RTV Drenthe uploadt interessant, het heet Arkiver. Dit account uploadt live webstreams en webvideo's, van over heel de wereld.¹⁰¹ Hier lijkt geen selectie in te zetten, op de dag dat deze alinea geschreven wordt (23 mei 2017) is er wederom een uur RTV Drenthe verschenen van deze bot.¹⁰² Het enige wat deze bot doet is vaste webstreams afgaan en daar een opname van maken.

De in deze paragraaf besproken artikelen zijn voornamelijk bronnen die worden bewaard als extra binnen The Internet Archive. Ze zijn al beschikbaar binnen andere archieven of worden uitgezonden. Hiermee wordt hun context in The Internet Archive vooral een soort van externe opslagplaats. Ze zijn beschikbaar in hun originele digitale archief en hier. Dit is voornamelijk vanuit The Internet Archive gedaan. Het nadelige hiervan is dat het niet wordt bijgehouden, waardoor de functie van externe opslagplaats niet blijvend is voor alle openbare archiefstukken uit een ander archief.

5.3 De Persoonlijke Gebruikersarchieven

Als de bots, Google Books en grote hoeveelheid bronnen van de KNB en Beeld en Geluid worden weggehaald, blijven de door de gebruikers aangelegde archieven over. Vanuit deze archieven komt de grote diversiteit binnen The Internet Archive naar voren. Tussen deze gebruikersarchieven zijn een aantal overeenkomsten. Ten eerste is er bij gebruikersaccounts vaak een eenmalige actieve upload van bronnen, daarna is het account niet meer actief geweest. Ten tweede hebben de meeste gebruikersaccounts een doel achter hun archief,

¹⁰⁰ <https://archive.org/search.php?query=Nederland&sort=-publicdate> (6 juni 2017).

¹⁰¹ <https://archive.org/details/@arkiver> (6 juni 2017).

¹⁰² TIA, RTV Drenthe 2017-05-23 09:19, https://archive.org/details/archiveteam_videobot_RTVDrenthe_20170523_0919 (6 juni 2017).

alhoewel er enkele accounts zijn met maar één upload, hebben accounts met meerdere uploads vaak hetzelfde thema of (politiek/maatschappelijke) doel.

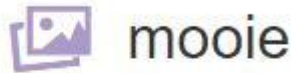
Voordat er wordt ingegaan op die doelen, een kleine beschrijving van een gemiddeld gebruikersaccount. Veel van de gebruikersaccounts hebben een anonieme gebruikersnaam, waardoor niet direct achterhaald kan worden wie daar achter zit. Het is niet mogelijk om direct contact op te nemen met deze personen. Een gebruikersaccount bestaat uit drie of vier tabbladen: Uploads, posts, collections en loans. Bij uploads staan alle stukken die de gebruiker aan het archief heeft toegevoegd, bij posts alle forumbijdragen op het forum van The Internet Archive, collections toont de door de gebruiker samengestelde collecties (een favoriete archiefstukken van [gebruiker]-collectie bijvoorbeeld) en loans toont de boeken uit The Internet Archive die iemand tijdelijk kan inzien/kan lenen.

Wat voor gebruikersarchieven zijn er te vinden? Om te beginnen met voorbeelden van kort-actieve accounts. Hans Steen (vermoedelijk een echte naam) heeft in een korte periode van 4 maanden (december 2006 tot maart 2007) een aantal video's en afbeeldingen in het archief gezet, 82 in totaal. De andere tabbladen zijn leeg.¹⁰³ De video's zijn allemaal amateuropnames, voornamelijk van ritjes en feestelijkheden. De afbeeldingen zijn waarschijnlijk door meneer Steen gemaakte foto's. Veel video's en foto's hebben abstracte namen als 'mooie', 'maatje' en 'hoekje'.

Om het archiefstuk 'mooie' iets dieper te analyseren.¹⁰⁴ Het is een foto van een boom die in bloei staat, met op de achtergrond een huis. De begeleidende tekst van Hans Steen is 'een mooie boom zomaar in de straat', het is het laatste archiefstuk wat door meneer Steen op zijn account is gezet. Voor een onderzoekend historicus of professionele archivaris mist hier een gigantische hoop context en dat geldt voor bijna het gehele archief van meneer Steen. Welke boom in welke straat? Wanneer is deze foto genomen, alleen de archiveringsdatum wordt vermeld? Dit lijkt een persoonlijk archief te zijn van meneer Steen, ondanks het doel van The Internet Archive om daar niet een verzamelplaats voor te zijn. (Zie afbeelding 3)

¹⁰³Gebruikersaccount Hans Steen, https://archive.org/details/@hans_steen (6 juni 2017).

¹⁰⁴TIA, mooie, <https://archive.org/details/mooie> (6 juni 2017).



een mooie boom zomaar in de straat

Author hans
Date_created 2007-03-17
Mature_content No
Monochromatic No
Mediatype image
Publicdate 2007-03-19 03:32:17
Identifier mooie
Backup_location ia903600_32

Afbeelding 3, beschrijving van het object. De foto zelf is alleen te bereiken door door te klikken, er is geen preview beschikbaar op de pagina zelf, (screenshot door auteur 14 juni 2017).

Het archief van Hans Steen is exemplarisch voor een aantal gebruikersarchieven binnen The Internet Archive. Deze pagina's worden voornamelijk gebruikt als een soort persoonlijke opslagpagina of The Internet Archive als Cloud. Ik als gebruiker wil foto's opslaan buiten mijn eigen opslagmogelijkheden, ik sla ze op via The Internet Archive. Het is mogelijk om video's van The Internet Archive weer te kunnen geven (officiële term embedden) op andere websites. Dit verklaart waarschijnlijk enkele accounts waar reclamefilmmpjes op staan.¹⁰⁵ Zijn deze archiefstukken dan bedoeld als archiefstuk of dient dit meer een persoonlijk doel?

Dit persoonlijke doel kan promotie zijn in plaats van bewaren, zoals in het geval van twee gebruikers: een goochelaar en een DJ.¹⁰⁶ Dit zijn tweearchieven die persoonlijke creaties promoten (goochel-acts en muziektracks), zij het dat door de creatieve inhoud er een meer

¹⁰⁵ Gebruikersaccount tinova, <https://archive.org/details/@tinova> (6 juni 2017) en gebruikersaccount kimberlybmichalec, <https://archive.org/details/@kimberlybmichalec> (6 juni 2017).

¹⁰⁶ Gebruikersaccount Marzzz <https://archive.org/details/@marzzz> (6 juni 2017) en gebruikersaccount magicshow, <https://archive.org/details/@magicshow> (6 juni 2017).

promotioneel karakter achter zit. De goochelaar lijkt zijn goochelkunsten door video's te willen promoten. Het DJ-account heeft naast zijn eigen remixes oudere bronnen als een VARA-radioshow geüpload¹⁰⁷ en daar is een mix van eigen archiefstukken en onpersoonlijke archiefstukken te vinden. De DJ is lang actief geweest, zijn eerste archiefstuk is uit 2007 en zijn laatste uit 2015.

Dat het persoonlijke kan mengen met het publieke spreekt uit de gebruikerspagina van Joep Zander, een actieve gebruiker (laatste upload bij tijd van schrijven was 22 mei 2017).¹⁰⁸ Deze man zet zich in voor ouders wiens kind geen contact met hen wil. Dit lijkt gedreven te zijn uit persoonlijke ervaringen en heel veel van zijn archiefstukken zijn publicaties van zijn hand over deze problematiek. Hier is het duale terug te zien van een persoonlijke en publieke functie van zijn archief. Enerzijds is dit een zelfgemaakte verzameling van zelfgeschreven publicaties, anderzijds toont het een gespecificeerde verzameling literatuur en de publieke acties van één man om een bepaald probleem binnen de maatschappij onder de aandacht te brengen. Dit is een archief waarin een publiekelijk doel en een persoonlijk doel sterk zijn vervlochten met elkaar.

De besproken gebruikersarchieven tot nu toe hebben allemaal een grotendeels persoonlijk karakter. Ze promoten allen op een manier de persoon die het archief heeft aangemaakt. Dit promoten heeft verschillende vormen, het archief is een opslag van gemaakte foto's of een promotieplek van het werk of doel van iemand. Opvallend is dat veel van dezearchieven maar korte tijd actief zijn gebruikt. Het zou kunnen dat The Internet Archive vooral is gebruikt als een soort alternatieve videospeler, in navolging van YouTube. Dit verandert de inhoud van deze gebruikersarchieven. Wordt The Internet Archive hier meer gebruikt als middel, in de zin dat het een videospeler biedt of een opslagplek dan als een archief waar een cultureel geheugen wordt geschapen?

5.4 De actieve gebruikers

Er zijn twee gebruikers die al lange tijd actief zijn, veel stukken op The Internet Archive hebben gezet en daardoor erg beeldbepalend zijn voor het archief. Ze zijn al redelijk beeldbepalend binnen de zoekterm Nederland, maar vooral binnen de kleinere zoektermen.

¹⁰⁷ TIA, De Avonturen van Ome Keesie [VARA 26 febr 1941]

<https://archive.org/details/DeAvonturenVanOmeKeesievara26Febr1941> (6 juni 2017).

¹⁰⁸ Gebruikersaccount Joep Zander, <https://archive.org/details/@joepzander> (6 juni 2017).

Het zijn twee gebruikers die niet een directe persoonlijke connectie hebben met de archiefstukken, al lijkt er bij beide een persoonlijke of grotere motivatie tot uploaden achter te zitten.

Als eerste is daar de gebruiker met de naam Filmcollective Worldwide.¹⁰⁹ De naam suggereert al een bepaalde internationaliteit. Dit account heeft nu een archiefcollectie opgebouwd van ruim negenduizend stukken, waarvan bij tijd van schrijven de laatste van 23 april 2017 is. Het is niet te achterhalen wie er specifiek achter dit account zit. Wat wel valt te deduceren is dat de persoon achter het account is aangesloten bij Genootschap Oud Zandvoort, een (amateur)vereniging die de geschiedenis van Zandvoort promoot en onderzoekt.¹¹⁰ Dit blijkt vooral uit het feit dat veel van de bronnen gearhiveerd zijn met de startletters GOZ (Genootschap Oud Zandvoort) en inhoudelijk van de gemeente Zandvoort of een ander aan Zandvoort verbonden bedrijf/instituut komen.¹¹¹ Deze bronnen bestaan voornamelijk uit de structuur GOZ-PDF-[nummer] en hebben geen duidelijke naam of uitgebreide metadata. Er staan op de website van GOZ een aantal links naar digitale kranten in The Internet Archive en die komen bij hetzelfde account uit. Dit bevestigt de verbondenheid.¹¹²

Daarnaast heeft Filmcollective Worldwide een aantal filmpjes met in de titel filmcollectief, met wisselende en ogenschijnlijk willekeurige inhoud. Er zijn filmpjes die in India te lijken zijn opgenomen en op boorplatformen in een heel ander deel van de wereld.¹¹³ Deze archiefstukken kennen dezelfde beschrijving in de vorm filmcollectief-SR-[nummer]. Ook hier ontbreken extra metadata over de meest recente stukken. Al deze video's zijn samengebracht in één verzameling binnen The Internet Archive: Collectie Filmcollectief.¹¹⁴ Deze collectie bevat meer dan vijfduizend stukken en lijkt volledig te zijn opgebouwd uit de films van Filmcollective Worldwide. Opmerkelijk is de wisselwerking tussen de woorden collectief en collective, anders dan de naam van het account wordt vooral de Nederlandse term gehanteerd. Veel van de films zijn trouwens afkomstig uit de regio Zandvoort. (Zie afbeelding 4).

¹⁰⁹ Gebruikersaccount Filmcollective Worldwide, https://archive.org/details/@filmcollective_worldwide (6 juni 2017).

¹¹⁰ Website Genootschap Oud Zandvoort, <https://www.oudzandvoort.nl/> (6 juni 2017).

¹¹¹ Bijvoorbeeld TIA, GOZ-PDF-H02181 <https://archive.org/details/GOZ-PDF-H02181> (6 juni 2017).

¹¹² Genootschap Oud Zandvoort, De Zandvoortsche Courant 2014

<https://www.oudzandvoort.nl/geschiedenis/digitale-kranten/572-de-zandvoortse-courant-2014> (6 juni 2017).

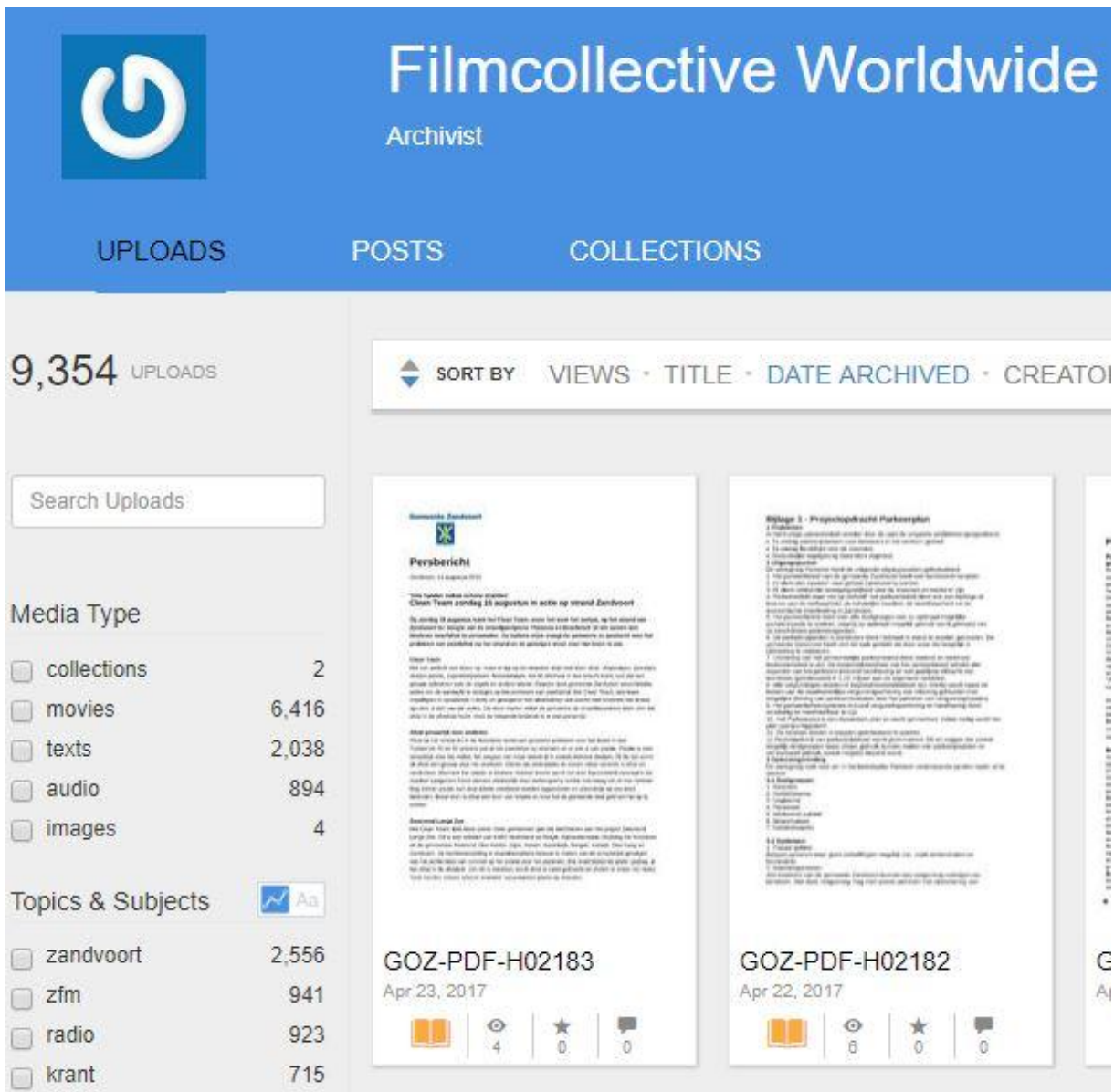
¹¹³ TIA, filmcollectief-SR-0067, <https://archive.org/details/filmcollectief-SR-0067> (6 juni 2017) en TIA, filmcollectief-SR-0072 <https://archive.org/details/filmcollectief-SR-0072> (6 juni 2017).

¹¹⁴ Collectie Filmcollective Worldwide, https://archive.org/details/collectie_filmcollectief?&sort=-publicdate (6 juni 2017).

Waarom is dit account in verhouding zo relevant? Door de grote omvang van het archief en de focus op Zandvoort komt Zandvoort binnen de kleine zoektermen met regelmaat terug. De zoekterm Tweede Wereldoorlog heeft een gedeeltelijk Zandvoorts tintje, hetzelfde geldt voor Nederlands-Indië, waar een oude video staat van een Zandvoortse familie.¹¹⁵ Bij de algemene zoekterm Nederland springt dit archief er vooral uit, omdat er een terugkerend thema over Zandvoort is terug te vinden. Er is zeker voor de zoektermen Tweede Wereldoorlog en Nederlands-Indië materiaal wat anders is dan de Zandvoortse context. Ondanks de focus op Zandvoort is de rest van het archief van Filmcollective Worldwide vrij breed opgezet. Zeker de filmpjes kennen een vrij brede oriëntatie en lijken veel verschillende makers te kennen, al worden die niet altijd genoemd. De codes als GOZ-PDF-[nummer] lijken trouwens later te worden aangepast en uitgebreid tot meer nuttige vindbare titels en beschrijvingen.¹¹⁶

¹¹⁵ TIA, Collectie Filmcollectief – Familie Koper in Nederlands-Indië <https://archive.org/details/familie-koper-uit-nederlands-indie> (6 juni 2017).

¹¹⁶ https://archive.org/details/@filmcollective_worldwide (6 juni 2017) in het overzicht valt op dat er bij oudere archiefstukken wel een volledige titelbeschrijving is.

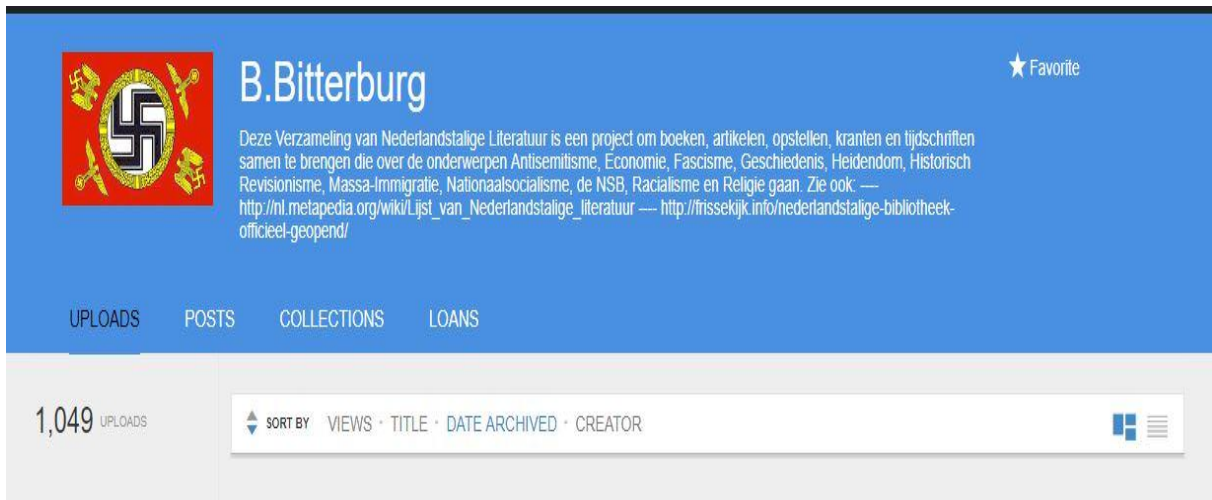


Afbeelding 4, Eerste resultaten gebruikersaccount Filmcollective Worldwide. Let op de nummering van de recentste twee bronnen. Ook is onder Topics & Subjects te zien dat archiefstukken met ‘zandvoort’ als onderwerp 2.556 keer in het archief terugkomen. (Screenshot door auteur 14 juni 2017).

Het is opvallend om te merken dat Zandvoort met regelmaat terugkomt binnen de zoekresultaten. Een opvallend patroon binnen dit onderzoek en er valt de vraag te stellen of bij specifiek onderzoek naar één onderwerp Filmcollective Worldwide zo opvalt in het geheel. Iemand die tevens opvalt in het grote geheel van onderwerpen is het gebruikersaccount van B. Bitterburg.¹¹⁷ Dit is een account met nu iets meer dan duizend uploads, de laatste bij tijd van schrijven uit februari 2017. Dit is het account wat het meest duidelijk zijn doel als archief beschrijft, is ook de enige zover gevonden die zijn gebruikersaccount naar zijn eigen doel

¹¹⁷ Gebruikersaccount Bitterburg, https://archive.org/details/@b_bitterburg (6 juni 2017).

heeft aangepast. Dit account heeft een beschrijvende tekst over zijn collectie bovenin en een archiefafbeelding (zie afbeelding 5). Wat opvalt aan deze archiefafbeelding is dat het een swastika betreft omringt met andere nazisymbolen zoals de adelaar.



Afbeelding 5, Uitleg over en afbeelding van het Bitterburg-archief in The Internet Archive. Let ook op aantal uploads linksonder (screenshot door auteur 6 juni 2017).

Wat goed beschreven is op dit account is het doel van het archief. Dit account beoogt een verzameling Nederlandstalige literatuur over een tal van onderwerpen, waar vooral fascisme, nationaalsocialisme, racialisme en antisemitisme uitspringen helemaal in een context van historisch revisionisme, ook één van de onderwerpen. Het account wil die literatuur bewaren voor later gebruik en niet in de vergetelheid brengen. Ook op de eigen website frissekijk.info is het doel van dit archief toegelicht.¹¹⁸ Deze site is tijdens het schrijven van dit onderzoek offline gegaan, maar terug te halen via de WaybackMachine.¹¹⁹ De exacte reden van het offline gaan van de website is voor de auteur van dit onderzoek niet duidelijk.

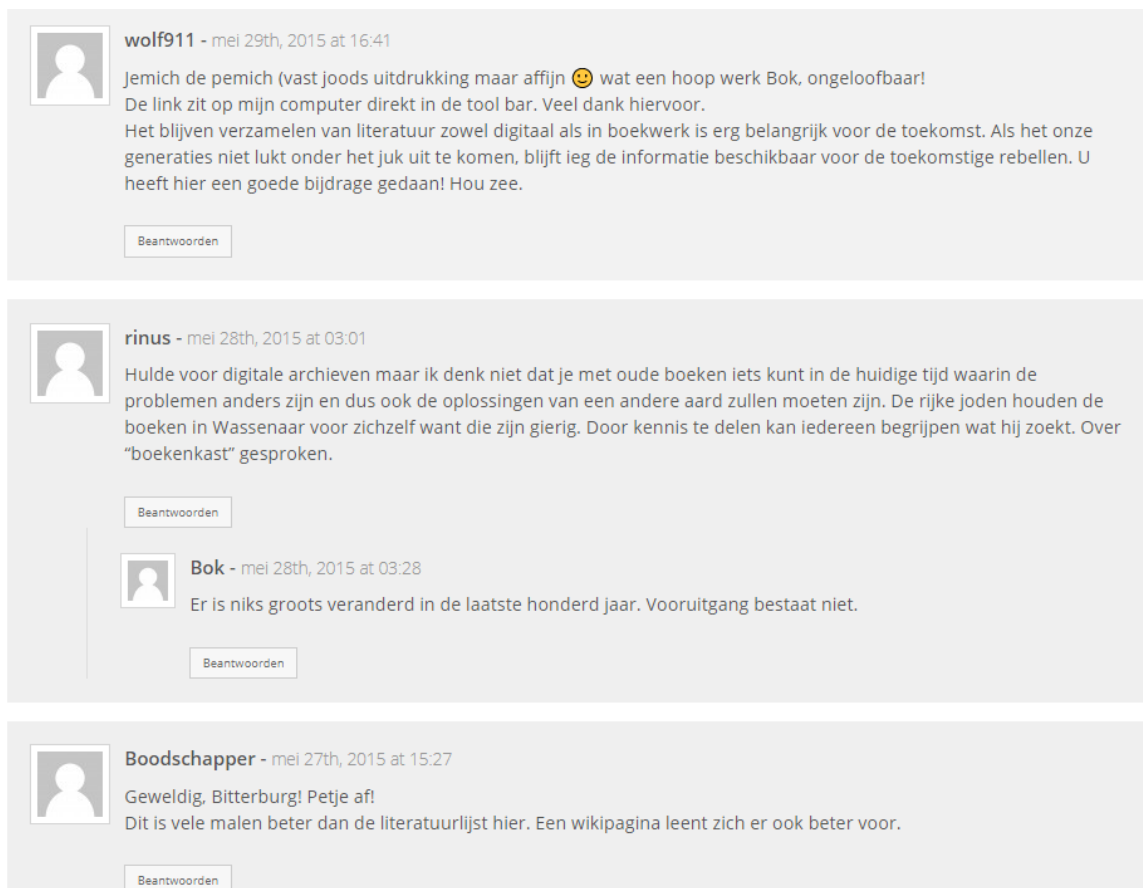
Op die website staat een nieuwsbericht uit 2015 over het aanleggen van dit archief. De auteur, hier onder de naam Bok, geeft enkele redenen waarom The Internet Archive zijn voorkeur heeft. Vooral de betrouwbaarheid van het instituut was voor hem belangrijk: onafhankelijk, bestaat al langere tijd, wordt steeds makkelijker doorzoekbaar en heeft

¹¹⁸ <http://frissekijk.info/nederlandstalige-bibliotheek-officieel-geopend/> (niet werkend 6 juni 2017).

¹¹⁹ Bok, 'Nederlandstalige Bibliotheek Officieel Geopend', (versie 27 mei 2015) <https://web.archive.org/web/20170106121821/http://frissekijk.info/nederlandstalige-bibliotheek-officieel-geopend/> (versie WaybackMachine 6 januari 2017) (6 juni 2017).

verschillende back-ups over de wereld. Een andere reden is het feit dat er al veel nazistische en revisionistische literatuur op The Internet Archive staan. Daarnaast heeft The Internet Archive geen ideologische achtergrond waar Bitterburg/Bok heel erg op tegen is.¹²⁰

Er zijn een zevental reacties op het artikel op zijn/hun website, allemaal geschreven een paar dagen na het online zetten van dit artikel in 2015, de weergave van The WaybackMachine is uit januari 2017. De meeste van deze reacties prijzen het werk of geven een suggestie tot aanvulling van het archief. Een ander suggereert dat het een goed initiatief is, maar dat alle rijke Joden de goede boeken achter houden. (Zie afbeelding zes).



The screenshot shows three comments on a website. Each comment includes a profile picture icon, the user's name, the date and time, and the text of the comment. Below each comment is a 'Beantwoorden' (Reply) button.

wolf911 - mei 29th, 2015 at 16:41
Jemich de pemich (vast joods uitdrukking maar affijn 😊) wat een hoop werk Bok, ongeloofbaar!
De link zit op mijn computer direkt in de tool bar. Veel dank hiervoor.
Het blijven verzamelen van literatuur zowel digitaal als in boekwerk is erg belangrijk voor de toekomst. Als het onze generaties niet lukt onder het juk uit te komen, blijft ieg de informatie beschikbaar voor de toekomstige rebellen. U heeft hier een goede bijdrage gedaan! Hou zee.

rinus - mei 28th, 2015 at 03:01
Hulde voor digitale archieven maar ik denk niet dat je met oude boeken iets kunt in de huidige tijd waarin de problemen anders zijn en dus ook de oplossingen van een andere aard zullen moeten zijn. De rijke joden houden de boeken in Wassenaar voor zichzelf want die zijn gierig. Door kennis te delen kan iedereen begrijpen wat hij zoekt. Over "boekenkast" gesproken.

Bok - mei 28th, 2015 at 03:28
Er is niks groots veranderd in de laatste honderd jaar. Vooruitgang bestaat niet.

Boodschapper - mei 27th, 2015 at 15:27
Geweldig, Bitterburg! Petje af!
Dit is vele malen beter dan de literatuurlijst hier. Een wikipagina leent zich er ook beter voor.

Afbeelding 6, reacties op artikel op frissekijk, via The WaybackMachine. Bok is hier de beheerder van het account Bitterburg. (Screenshot door auteur 14 juni 2017).

Dit laatste voorbeeld om kort te illustreren wat de inhoud is van het archief. Uit alle berichten van de gebruiker zelf en dergelijke spreekt dat het hoogstwaarschijnlijk een persoon is met een neonazistische wereldvisie. Zeker de eerste grote upload rond 2015 bevatten voornamelijk NSB-documentatie en boeken, allen met een sterk pro-nationaalsocialistische

¹²⁰ Bok, 'Nederlandstalige Bibliotheek Officieel Geopend', (versie 27 mei 2015) <https://web.archive.org/web/20170106121821/http://frissekijk.info/nederlandstalige-bibliotheek-officieel-geopend/> (versie WaybackMachine 6 januari 2017) (6 juni 2017).

inhoud. De collectie bestaat hoofdzakelijk uit boeken, deze zijn soms gereproduceerd van andere websites, zoals wordt aangegeven op frissekijk.info en een aantal video's.

De meest recente archiefstukken hebben allemaal betrekking tot de rechtse politicus Hans Janmaat.¹²¹ Hier moet bij worden gezegd dat niet alle stukken sterk ideologisch zijn. In de Hans Janmaat-stukken zitten bijvoorbeeld een uitzending van het televisieprogramma *Andere Tijden* en een satirische benadering van Hans Janmaat door het programma *Keek op de Week van de Nederlandse satirici Van Kooten en De Bie*.¹²² Opmerkelijk is hier dat de *Andere Tijden* aflevering niet als zodanig is genoemd. Het archiefstuk heeft de titel *Wilders Is Vele Malen Erger Dan Janmaat* en heeft als verdere beschrijving alleen een korte biografie van Janmaat.

Er is bijvoorbeeld een rapport uit 2013 van het Sociaal Cultureel Planbureau te vinden over integratie.¹²³ In de context van het totale gebruikersactief zal daar misschien een boodschap achter zitten. Als dit rapport bij een willekeurige zoektocht gevonden wordt, zit er niet de context van dat gebruikersarchief bij, maar is het gewoon een onderzoeksrapport over integratie in Nederland, zonder de fascistisch en nazistische context van het gebruikersarchief. Curieus genoeg is er ook een boek over eetbare planten en paddenstoelen in Nederland te vinden in het Bitterburg-archief.¹²⁴

Voor Bitterburg geldt hetzelfde als voor het Zandvoortarchief, door de grote hoeveelheid en activiteit, in dit geval door bijna alleen maar Nederlandstalige literatuur en video's, drukt dit archief heel erg een stempel op de zoektermen. Een enkele bron over Zandvoort valt in het grote geheel wel mee. Bitterburg, zeker bij de controversiële zoekonderwerpen, komt regelmatig naar voren. Bij de zoekresultaten over de Tweede Wereldoorlog komen veel stukken van hem naar voren. Door de ideologische inhoud kleurt dit heel erg de gevonden stukken. Er wordt niet automatisch tegengewicht geboden aan deze gekleurde inhoud. Bitterburg heeft vrij spel om zijn visie dominant te maken binnen het archief. Zeker zijn controversiële inhoud vraagt om andere archiefstukken die dit nuanceren.

¹²¹ https://archive.org/details/%40b_bitterburg?&sort=-publicdate (6 juni 2017).

¹²² TIA, *Wilders is Vele Malen Erger dan Janmaat* <https://archive.org/details/WildersIsVeleMalenErgerDanJanmaat> (6 juni 2017) en TIA, *Keek op de Week – 96 = 3 van 7 – Centrum Democraten* <https://archive.org/details/KeekOpDeWeek963van7CentrumDemocraten> (6 juni 2017).

¹²³ TIA, *Jaarrapport Integratie 2013* <https://archive.org/details/JaarrapportIntegratie2013> (6 juni 2017).

¹²⁴ TIA, *Nederlandse Eetbare Planten en Paddenstoelen Database* <https://archive.org/details/NederlandseEetbarePlantenEnPaddenstoelenDatabase> (6 juni 2017).

5.5 Losse archiefstukken

Naast actieve gebruikers met een breder of persoonlijker archief zijn er een aantal door gebruikers geüploade losse archiefstukken te vinden. Vaak hebben die gebruikers maar één archiefstuk, waarbij het heel moeilijk is om te kunnen achterhalen wat de originele motivatie was om dit archiefstuk in The Internet Archive te zetten. Wat vaak ontbreekt bij deze archiefstukken is context. Ze hebben een onduidelijke titel en onduidelijke of vaak ontbrekende omschrijving.

Er zijn een aantal voorbeelden van, deze staan in de literatuurlijst. Hier bespreken we er enkelen. Een goed voorbeeld van vage context is het archiefstuk getiteld ‘Konijnntje Official’.¹²⁵ Dit is geüpload door de gebruiker ‘Rainbow Fantasy’ en is het enige archiefstuk wat hij/zij heeft geüpload. Het is een grote afbeelding van een webpagina Tumblr. Er zijn geen aanduidingen dat het een screenshot is, maar het lijkt een zelf samengevoegde afbeelding, waarop enkele strips van het stripfiguur Konijn staan. Voor de rest is er weinig te zien, alleen dat deze strips in 2013 zijn gemaakt. Dit is weer een grotendeels persoonlijk archiefstuk wat zonder context of duidelijke motivatie wordt gepresenteerd.

Een ander voorbeeld is vooral een voorbeeld van verkeerd archiveren, het archiefstuk ‘Vriendjespolitiek’.¹²⁶ De bron wordt beschreven als een uitzending van het NOS-programma EenVandaag van 9 februari 2010. Echter is er alleen een gedigitaliseerde bladzijde te zien met daarop het logo van de Socialistische Partij (SP), waar SP Studio staat. SP Studio is het videoprogramma van de SP waar ze mensen interviewen.¹²⁷ Vervolgens staat er alleen in het document de tekst ‘SP (vriendjes politiek)’ en na een aantal regels wit een weblink naar de website van de SP. Deze link werkt niet en krijgt bij tijd van schrijven een foutmelding van de SP-website.¹²⁸ Met een zoektocht op die website, mede door hulp van de in de foutieve link beschreven woorden, is het waarschijnlijke originele artikel toch te achterhalen.¹²⁹ De datum komt overeen en er wordt in dit artikel verwezen naar een EenVandaag uitzending, die niet meer te bekijken is. (Zie afbeelding 7)

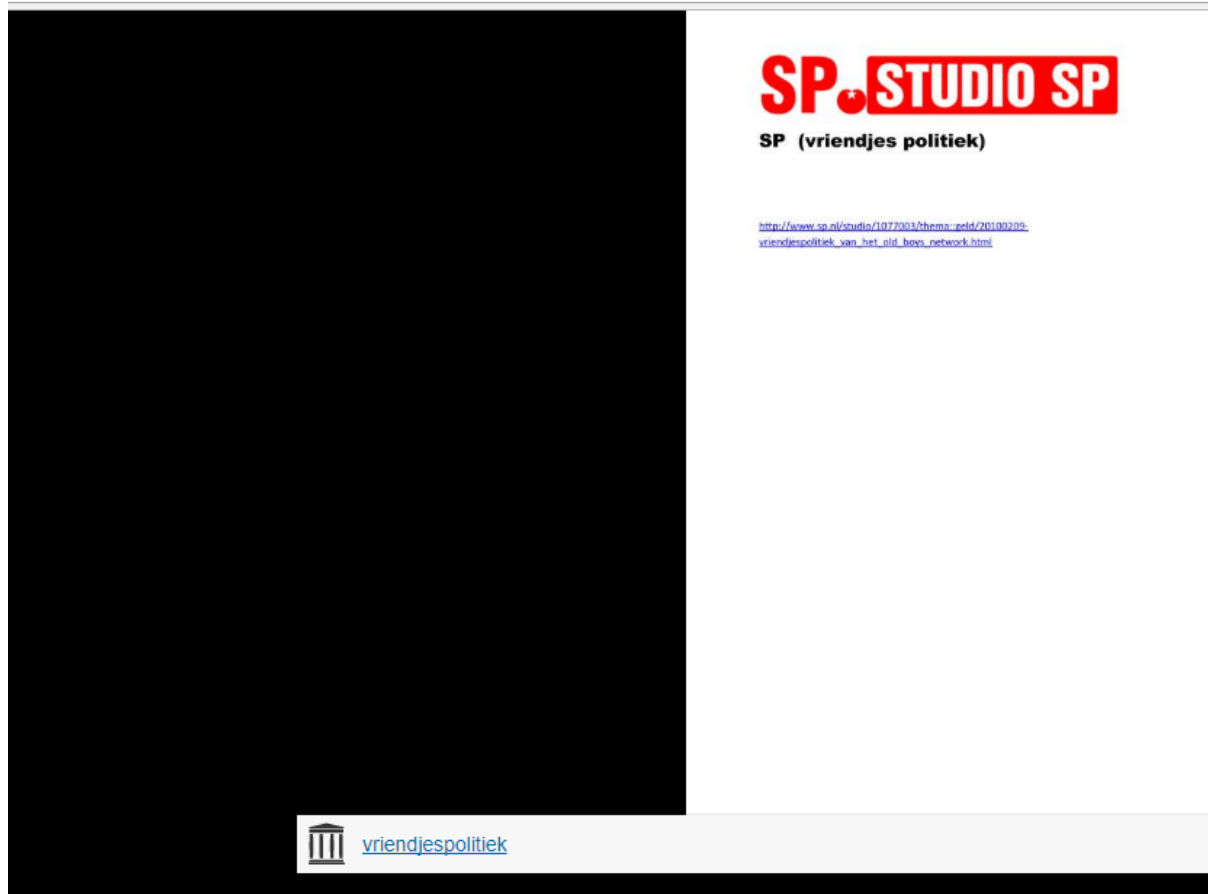
¹²⁵ TIA, Konijnntje Official <https://archive.org/details/KonijnntjeOfficial> (6 juni 2017).

¹²⁶ TIA, vriendjespolitiek <https://archive.org/details/vriendjespolitiek> (6 juni 2017).

¹²⁷ SP, ‘Emile Roemer: ‘Er is genoeg voor iedereen’’, <https://www.sp.nl/video/emile-roemer-er-is-genoege-voor-iedereen> (6 juni 2017).

¹²⁸ <https://www.sp.nl/studio/1077003/thema::geld/20100209-vriendjespolitiek-van-het-old-boys-network.html> (foutmelding 6 juni 2017).

¹²⁹ SP, ‘Bedrijfsleven lobbyt tegen beperking aantal nevenfuncties door SP’ (versie 6 februari 2010) <https://www.sp.nl/nieuws/2010/02/bedrijfsleven-lobbyt-tegen-beperking-aantal-nevenfuncties-door-sp> (6 juni 2017).



 vriendjespolitiek
by NOS 2010

Usage [Public Domain Mark 1.0](#)  
Topics [education](#) and [non-commercial](#)

Afbeelding 7, weergave van het document 'vriendjespolitiek', dit is geen televisie-uitzending. EenVandaag is niet van de NOS, toch wordt dat hier als auteur of maker weergegeven. (Screenshot door auteur 14 juni 2017).

Er worden bij dit originele archiefstuk in The Internet Archive een aantal dingen foutief weergegeven. Deze bron is geen uitzending van EenVandaag, maar een document waar een link staat die verwijst naar een artikel over een uitzending van EenVandaag. Verder wordt EenVandaag beschreven als een NOS-programma, EenVandaag wordt al sinds 2004 en sinds 1993 uitgezonden door respectievelijk de AVRO en de TROS-omroepen, het huidige AVROTROS.¹³⁰ Het archiefstuk zelf is een leeg blad met een logo en weblink, waarvan de weblink geen resultaat geeft. Dit stuk heeft geen nut meer binnen The Internet Archive.

¹³⁰ <http://www.eenvandaag.nl/> (6 juni 2017).

Er zijn een aantal gebruikers en losse bronnen die in metadata soms te kort schieten of onduidelijk/foutief zijn. Wegens de ruimte binnen dit onderzoek is er gekozen om de meest exemplarische voorbeelden te noemen, veel van dezelfde trends zoals hierboven besproken zijn terug te vinden bij de andere gebruikers. In de literatuurlijst is een uitgebreidere lijst van deze gebruikers en bronnen opgenomen. Ook bij deze bronnen is het opmerkelijk dat er weinig tot geen controle is op dit soort uploads. Er wordt weinig aan gedaan om stukken vindbaar te maken of om fouten te corrigeren.

5.6 Hoofdstukconclusie

Wat is er opvallend binnen het Nederlandse gedeelte van The Internet Archive? De meeste bronnen en gebruikers voegen archiefstukken toe met een persoonlijke of politieke motivatie. Die motivaties zijn divers, waardoor het archief ook erg divers is. Gebruikers die heel actief zijn, zijn sneller terug te vinden, waardoor hun motivatie, zeker bij de kleinere gehanteerde zoektermen, de overhand kan nemen binnen de archiefstukken. Veel archiefstukken hebben foutieve of incomplete metadata. Beschrijvingen en titels van stukken zijn vaak onduidelijk, niet compleet of foutief.

Wat vooral sterk opvalt is het weinige tegengewicht en kritiek die deze bronnen krijgen. Het nazistische archief is actief bezig om hun visie op de geschiedenis te promoten via The Internet Archive enerzijds en hun eigen websites anderzijds. Wat zeker bij de kleinere zoektermen opvalt is dat er weinig bronnen zijn die deze activiteit tegen kunnen gaan. In principe moet er nu een andere gebruiker of historicus opstaan die bronnen toevoegt om de alternatieve visie van het Bitterburg-archief tegen te gaan. Alhoewel de andere bronnen minder extreem zijn, geldt daar hetzelfde voor. Het is opmerkelijk dat er slecht beschreven bronnen in het archief staan en dat daar niet op lijkt te worden gecontroleerd. Dit wordt versterkt doordat bronnen niet direct zijn aan te passen en gebruikersaccounts niet direct zijn te contacteren. Wat dit betekent voor de status van The Internet Archive voor historisch onderzoek wordt in het volgende hoofdstuk besproken.

Hoofdstuk 6 The Internet Archive voor historisch onderzoek

Dit hoofdstuk wil een antwoord vinden op de hoofdvraag van dit onderzoek: Is het nuttig om historisch onderzoek te doen naar The Internet Archive? In hoofdstuk 4 is besproken dat veel onderzoek naar The Internet Archive nu vooral hun website-archief betreft of de technische middelen waarmee ze de websites verzamelen. Het enige inhoudelijke onderzoek van Anya Boyle argumenteerde dat het door haar bekeken archief te ongeordend was en een te persoonlijke aard had. Dit hoofdstuk bekijkt hetzelfde maar dan aan de hand van de bredere gebruikersarchieven uit hoofdstuk 5.

Is de inhoud van The Internet Archive nuttig om historisch onderzoek naar te doen? Dit onderzoek twijfelt daar zeer aan en zal in dit hoofdstuk betogen waarom. Dit wordt gedaan in gedeeltes. In het eerste gedeelte wordt de archieftheorie uit hoofdstuk 2 teruggehaald, van belang binnen archieven waren ordening en selectie met als recentste postmoderne visie: het creëren van een cultureel geheugen in samenspraak met de gebruikers. In het eerste gedeelte van het betoog wordt besproken hoe deze begrippen gehanteerd worden in The Internet Archive.

In het tweede gedeelte wordt een historisch onderzoek in The Internet Archive theoretisch gedaan. Vooral de start van een historisch onderzoek staat hier centraal. Welke bronnen zijn er te vinden? Wat zeggen die bronnen en vooral hoe nuttig zijn die bronnen in zijn totaliteit? Hier worden vooral de in hoofdstuk 5 besproken resultaten vergeleken en besproken wat ze kunnen betekenen voor de historicus. In het derde en afsluitende gedeelte wordt besproken hoe The Internet Archive beter gemaakt zou kunnen worden om historisch onderzoek te doen.

6.1 Ordening, selectie en een cultureel geheugen in The Internet Archive

Hoe past een digitaal deels door gebruikers gedreven archief als The Internet Archive binnen de discussies rondom archieven en digitale archieven? Deze discussies zijn besproken in hoofdstuk 2. Een korte samenvatting van deze discussie is dat het binnen het professionele archief er twee kernbegrippen belangrijk zijn: ordening en selectie. Ordening om een goed doorzoekbaar en zichtbaar archief te hebben. Selectie om de overmaat aan potentiële archiefstukken en ruimtetekort te beperken en om sterk na te denken over de relevantie van bronnen voor de toekomst. Beiden zijn door het professionele digitale archief en het

postmodernisme veranderd. Ordening is door zoektermen meer flexibel en fluïde geworden. Selectie wordt niet meer alleen gedaan door de professional, maar ook door de gebruikers. De onbeperkte digitale opslag maakt het mogelijk dat veel meer bewaard kan worden, waardoor selectie minder nodig kan zijn, al zijn daar twijfels over. Het postmoderne doel van een archief is veranderd van een semi-gesloten bewaarplaats naar een open instituut wat met zijn bronnen een cultureel geheugen probeert te scheppen. Hierbij is de invloed van externe factoren erg belangrijk en hebben gebruikers een centrale plaats.

Hoe zit het met de ordening en selectie binnen The Internet Archive? De ordening is heel erg flexibel. Op de startpagina van het archief zijn de verschillende onderdelen te vinden, allemaal heel erg breed ingedeeld in categorieën als websites, video's, audio en afbeeldingen. Opmerkelijk is dat er een onderscheid wordt gemaakt tussen video en televisieprogramma's, terwijl de laatste natuurlijk onder de noemer video valt. Het is niet helemaal duidelijk of onder video met ruime 3,3 miljoen resultaten ook het televisie-archief valt met 1,4 miljoen resultaten. Hetzelfde geldt voor het kopje audio en het kopje concerts (waar ook weer video-opnames onder kunnen vallen). In ieder geval omvatten al deze onderdelen minimaal honderdduizenden tot miljarden archiefstukken.

Onder de onderdelen staan op de startpagina populaire collecties die gemaakt zijn van deze archiefstukken, waaronder het Live Music Archive. In de top tien bij tijd van schrijven staat één collectie met een kleine vijfhonderd stukken, de kleinste collectie daarna bevat al meteen meer dan tienduizend stukken. Er is zeker sprake van ordening, echter is die ordening over het algemeen heel ruim en naar de mening van de auteur te algemeen. The Internet Archive is het beste te gebruiken door specifiek te zoeken naar onderwerpen, wat dit onderzoek heeft gedaan.

Echter bij het zoeken naar specifieke onderwerpen past The Internet Archive ook een bepaalde ordening toe. De zoekresultaten staan namelijk standaard geordend op de vage term 'Relevance' of relevantie. Het is niet helemaal duidelijk hoe die relevantie gemeten wordt. Er wordt gezegd hoeveel keer een individueel archiefstuk bekeken is, waarbij je zou kunnen aannemen dat een veelbekeken archiefstuk relevant zou zijn, want anderen hebben daarop geklikt. Het kan betekenen dat het woord of de woordgroep waarop wordt gezocht het meest voorkomt in die bepaalde bron. Een verhouding tussen de mensen die specifiek gezocht hebben op een term en procentueel binnen die term veel op één bron hebben geklikt, lijkt de meest logische verklaring.

Dit laatste is te illustreren met de zoekterm ‘Nederland’ en dan zoekend op relevantie.¹³¹ Bij tijd van schrijven komt eerst een aantal wedstrijden boven van het Nederlands voetbalelftal uit het wereldkampioenschap voetbal uit 2006. De eerste twee resultaten zijn voetbalwedstrijden, waarbij de eerste 188 keer is bekeken en de tweede 131 keer. In perspectief het zevende resultaat is ruim tweeduizend keer bekeken, het vijftiende resultaat ruim vierentwintigduizend keer. Het kan zo zijn dat mensen die hebben gezocht op ‘Nederland’ eerder op de voetbalwedstrijden hebben geklikt, dan op die andere meer bekeken resultaten. Waardoor percentueel gezien die sneller boven komen. Dit is speculatie.

Hoe zit het met selectie binnen The Internet Archive? Dit onderzoek durft te stellen dat daar bijna geen sprake van is. De boven besproken grootte is daar al een voorbeeld van. Miljoenen archiefstukken zitten er in het archief en daar lijkt geen selectie in te worden gemaakt. Dit valt op door de aanwezigheid van slecht beschreven bronnen geüpload door gebruikers. Bitterburg koos voor het gebruik van The Internet Archive, omdat er al veel fascistische literatuur op stond. Dit geeft aan dat er op de inhoud niet streng wordt geselecteerd en bronnen met een sterk ideologische inhoud niet worden geweerd uit het archief.

De enige selectie die binnen The Internet Archive genoemd kan worden is het onderscheid tussen openbare en leenbare stukken. De laatste zijn in veel mindere getalen aanwezig, van de ruim 8000 archiefstukken onder de zoekterm ‘Nederland’, kunnen er vier alleen op leenbasis worden ingezien. Dit heeft hoogstwaarschijnlijk met rechten te maken van die stukken en daarom wordt er een soort bibliotheekmodel op toegepast. Dit betekent vooral dat een boek niet door iedereen tegelijk bekeken kan worden, maar in beurten. Dit is het enige selectieve onderscheid tussen de miljoenen bronnen.

De ordening van The Internet Archive moet vooral zelf gemaakt worden. Doormiddel van het hanteren van zoektermen moet de grote hoeveelheid archiefstukken worden bedwongen, door de onderzoeker en door de reguliere gebruiker. The Internet Archive is wat dat betreft een uiterst flexibel archief. De keuze van The Internet Archive om nauwelijks te selecteren op bronnen is een opmerkelijke in verhouding met andere digitalearchieven. De postmodernistische invloed op de archieftheorie is vooral dat er een dialoog moet zijn over deze selectie en er doormiddel van die dialoog en de archiefstukken een cultureel geheugen

¹³¹ <https://archive.org/search.php?query=Nederland> (6 juni 2017).

kan worden geschapen. Slaat The Internet Archive door niet te selecteren een heel andere richting in dan het postmoderne debat?

Want is er sprake van een cultureel geheugen binnen The Internet Archive? Veel van de in hoofdstuk 5 besproken archieven hebben een sterk individueel karakter. Sommige daarvan zijn deels promotioneel (de goochelaar en de DJ), maar daar kan bij worden afgevraagd of ze The Internet Archive gebruiken als archief of als middel om hun werk te promoten. Er zijn veel losse stukken en archieven die één persoon (Joep Zander) of één plek (Zandvoort) promoten. In het meest controversiële geval wordt het archief gebruikt om de geschiedenis te herschrijven door het toevoegen van veel bronnen, dit is een doel van het Bitterburg-archief.

Een cultureel geheugen is in feite heel breed, er is in principe ruimte voor een neonazistische stem binnen het archief. Zeker de boeken uit het Bitterburg-archief die geschreven zijn ten tijde van de NSB en de nazi-bezetting hebben een zeker historische waarde. Er is geen stem binnen The Internet Archive die actief deze nazistische literatuur tegenspreekt. Het cultureel geheugen wordt niet aangevuld, simpelweg omdat er geen gebruiker is die de documenten van Bitterburg in een andere context plaatst.

De combinatie van weinig ordening en selectie en de hele individuele inhoud van het archief maken het zo dat er geen cultureel geheugen is. Dit wordt versterkt door het vrij laten van sterke stemmen zoals Bitterburg waar geen andere stemmen tegenover worden gezet. The Internet Archive fungeert vooral als een opslagplaats in veel gevallen, waarbij het niet of nauwelijks de bedoeling lijkt te zijn dat ze voor onderzoek gebruikt worden.

6.2 Onderzoek in The Internet Archive

Om het laatste gemaakte punt te illustreren een samenvatting van wat een historisch onderzoeker zou kunnen tegenkomen, mocht hij zijn onderzoek uitvoeren binnen The Internet Archive. Dit gekoppeld aan de theorie uit hoofdstuk 2 en 3 over professionele en persoonlijke digitale archieven. Het begin van het onderzoek wordt nu gesimuleerd, de verkenning van mogelijke bronnen. Wat gebeurt er als een onderzoeker dit zou doen binnen The Internet Archive? Onderstaande bevindingen gelden voor bijna alle zoekresultaten, maar voor het gemak en de relatieve grootte van enkele honderden resultaten is er voor de zoekterm ‘Tweede Wereldoorlog’ gekozen.

Wat voor bronnen komt de onderzoeker tegen? Er zitten inhoudelijk zeker interessante bronnen bij. Toch zijn veel van deze oudere bronnen afkomstig uit andere archieven die voor een Nederlandse onderzoeker niet onbekend zouden moeten zijn, onder andere Beeld en Geluid en de KNB. Voor de onderzoeker is het meteen wenselijker om die archieven te raadplegen, via hun digitale kanalen of het fysieke instituut. Waarom direct contact? De bronnen op The Internet Archive zijn bij tijd van schrijven van Beeld en Geluid al weer 6 jaar oud en van de KNB 3 jaar en de collecties van beide zijn niet door henzelf op The Internet Archive gezet. Direct contact kan meer boven brengen en niet-openbare bronnen tonen. Daarnaast komen alle Beeld en Geluid bronnen van hun openbeelden-project af, een project wat nog steeds actief is, maar deze zijn sinds 2011 niet meer op The Internet Archive gezet.

Daarnaast is men in Nederland bezig zoals in hoofdstuk 2 beschreven om een groter netwerk digitaal erfgoed mogelijk te maken. De digitale archieven van Nederland gaan samenwerken, om de digitale grenzen tussen verschillende archieven te overbruggen. Dit is met de strategie uit 2015 in gang gezet. Dit betekent het maken van een soortgelijk archief als The Internet Archive, want veel verschillende bronnen uit verschillende collecties. Echter zit er achter dit samenwerkingsverband veel meer expertise en zijn het altijd de beheerders van het eventuele fysieke gedigitaliseerde materiaal. Een onderzoeker heeft waarschijnlijk meer aan een bezoek aan een archief of overkoepelende website die al die expertise en achtergrondkennis waarborgt, dan aan een verzamelplaats als The Internet Archive.

Als de onderzoeker de bronnen van andere archieven buiten beschouwing laat, wat vindt hij nog meer in The Internet Archive? Veel door gebruikers geüploade bronnen, hier zijn twee problemen mee: kleuring en beschrijving. De archieven die het meest opvallen binnen alle zoektermen die dit onderzoek gehanteerd heeft, zijn toch de actieve gebruikersarchieven die één specifiek onderwerp voorop laten staan. In het geval van de Nederlandse bronnen zijn dit Zandvoort en de neonazistische literatuur van Bitterburg. Zeker voor de zoekterm ‘Tweede Wereldoorlog’ valt op dat er nogal wat gekleurde literatuur verschijnt. Dit kan ook in een professioneel digitaal archief natuurlijk, maar bij The Internet Archive zit daar geen bewuste verhouding in. Als opvalt dat alle sterk gekleurde bronnen te herleiden zijn naar één account, namelijk Bitterburg, dan geeft dat reden tot nadenken over de betrouwbaarheid van het totale archief. Hetzelfde geldt voor steeds terugkerende bronnen over Zandvoort. Er zijn zeker van andere gebruikers weinig bronnen binnen The Internet Archive die de grote hoeveelheid van Bitterburg en Filmcollective Worldwide in een ander daglicht zetten.

Er kan voor deze gebruikers worden gezegd dat The Internet Archive ruimte biedt voor onderbelichte bronnen. Bitterburg beschrijft op zijn website de noodzaak voor een verzameld archief met alternatieve literatuur en vindt de middelen om dit te bereiken op The Internet Archive. Als onderzoeker is het interessant om deze bronnen door te nemen, alleen worden geen nuances of achtergrond gegeven bij deze bronnen. Bitterburg hoeft geen verantwoording af te dragen voor het uploaden van deze bronnen en er is geen reflectie op de inhoud. Geen reflectie op de inhoud van het individuele stuk en geen reflectie op de totale inhoud van het Bitterburg-archief. Er wordt niet door andere gebruikers of The Internet Archive met andere bronnen tegengewicht gegeven aan de alternatieve historische visie die Bitterburg tentoon wil spreiden.

Als laatste vindt de historisch onderzoeker veel bronnen waarbij context ontbreekt. Bronnen die geen duidelijke titel dan wel beschrijving hebben. Bronnen die anders zijn dan doet vermoeden, denk hierbij aan de EenVandaag-bron uit het vorige hoofdstuk. Dit maakt deze bronnen voor onderzoek onbruikbaar, omdat de authenticiteit niet gecontroleerd kan worden. Dit geldt in zekere mate ook voor de bronnen die door gebruikers worden geüpload. Wie controleert hun metadata? Waar hebben zij deze stukken vandaan gehaald? Bitterburg haalt op zijn website aan dat hij in zijn gebruikersarchief verschillende stukken uit andere collecties heeft samengebracht. Welke collecties en hoe betrouwbaar zijn deze?

Er kunnen heel veel vragen aan het archiefmateriaal gesteld worden, maar de betrouwbaarheid is hetgeen wat het meest ontbreekt. Er is geen mogelijkheid om contact te zoeken met de gebruiker-accounts, je weet niet precies waar deze bronnen vandaan komen. Daarnaast is het aantoonbaar dat bronnen foutieve informatie bevatten, zoals de EenVandaag-bron. Er wordt achtergrondinformatie en andersoortige metadata meegegeven aan deze bronnen, het is alleen niet duidelijk hoe goed deze zijn gecontroleerd. Dit maakt bijna elke bron onbetrouwbaar binnen The Internet Archive.

6.3 Mogelijke oplossingen

De bronnen zijn onbetrouwbaar en er is geen cultureel geheugen, noch selectie, noch ordening. Dit maakt onderzoek doen naar de inhoud van The Internet Archive erg moeilijk. Hoe kan dit worden verbeterd? Er valt wat voor te zeggen dat selectie in een zo breed opgezet archief als dit niet of in mindere mate nodig is. De vrijheid die gebruikers hebben om hun eigen archiefstukken toe te voegen is nergens anders mogelijk en geeft The Internet Archive een interessant karakter. The Internet Archive is heel interessant om als publiekshistoricus te

bekijken, omdat veel mensen hun eigen interpretatie van het verleden laten zien of een persoonlijk archief achter laten.

Echter zou er een strengere controle op deze archiefstukken moeten plaatsvinden. Richtlijnen voor het goed beschrijven van een bron zou al een veel duidelijker archief geven. Dit hoeft niet eens per se door professionals gedaan te worden, dit kan ook door vrijwilligers/andere gebruikers. Deels biedt The Internet Archive de mogelijkheid al aan, er kunnen namelijk over archiefstukken reviews geplaatst worden door andere gebruikers, echter zeker bij kleine archiefstukken wordt dit niet gedaan, want deze zijn moeilijk te vinden. Er is niet een directe dialoog aan te gaan tussen de uploader en andere gebruikers. Het is wenselijk om, bijvoorbeeld net als Wikipedia, een discussiepagina aan een bron toe te voegen en richtlijnen op te stellen over hoe een goed beschreven bron, vooral in metadata eruit moet zien.

De verbetering van de ordening kan vooral gebeuren door het specificeren van de collecties. De collecties zijn nu heel breed opgezet of individueel samengesteld, dit kan specifieker. Ook door de losse onderwerpen en tags waarmee stukken worden samen gebracht te koppelen aan collecties. Iets met de zoekterm ‘Tweede Wereldoorlog’, moet in de collectie ‘Tweede Wereldoorlog’ komen. Er kan gezegd worden dat dit al geldt als je zoekt op ‘Tweede Wereldoorlog’, maar het zou handig zijn als de collectie juiste heldere bronnen weergeeft. Bitterburgs archief past dan binnen de collectie, maar wordt in een bredere context geplaatst. Dit vraagt natuurlijk wederom om een strengere beschouwing van bronnen. Een bron met historisch foutieve informatie heeft een plek, maar moet als zodanig geïdentificeerd worden.

Het cultureel geheugen heeft hier baat bij, collecties geven namelijk een lijn weer. Daarnaast zal het handig zijn als archiefstukken beter verbonden worden. Veel commerciële databases en archieven hebben een suggestie bij hun stukken. Dit kan zijn in de trant van ‘anderen bekeken ook:’ en dan enkele bronnen die veel bekeken zijn in relatie met de bron die nu wordt bekeken. Dit zorgt voor verbinding tussen de bronnen en de gebruikersarchieven binnen The Internet Archive en een bredere context. The Internet Archive heeft deze suggesties ook als er helemaal wordt doorgekeken naar onder aan de pagina van het archiefstuk. Echter zijn deze suggesties gebaseerd op metadata. Dit onderzoek concludeert dat veel van deze metadata niet volledig of juist is. De suggesties die The Internet Archive nu geeft komen vaak terug bij dezelfde gebruikersaccount als de originele bron. Bij een archief

als dat van Bitterburg versterkt dit weer zijn ideologische visie. Ook hier is controle op juiste metadata erg van belang.

Verder is het een aanrader om betrouwbaarheid van bronnen te versterken. Dit kan deels door de hierboven genoemde oplossingen, maar ook door professionals meer te betrekken bij het archief. Als in Nederland de laatste stap is dat digitale archieven veelal samen gaan werken, dan zou dat ook met The Internet Archive moeten kunnen. Het is vrij opmerkelijk dat een archivaris van The Internet Archive apart de open beelden van Beeld en Geluid online zet en deels storend dat dat niet systematisch wordt gedaan. Er valt veel te winnen als expertise en externe kennis bij de bronnen betrokken worden.

6.4 Hoofdstukconclusie

Is het nuttig om als historisch onderzoeker onderzoek te doen naar de inhoud van The Internet Archive? Inhoudelijk is dit niet de moeite waard. Er is een te ruime ordening en te weinig selectie om snel nuttige bronnen te zoeken. Alhoewel er collectief een archief verzameld wordt zit er geen lijn in het archief. Het archief voelt ondanks zijn grootte heel individueel aan, doordat er of automatisch lukraak verzameld wordt of door individuele gebruikers met elk een eigen doel. Door de ruime ordening is elke zoektocht binnen het archief een individuele. Een onderzoeker moet zelf tussen de vele bronnen zoeken naar bruikbare bronnen en heel erg veel rekening houden met de visie of onderwerp van één specifiek (gebruikers)archief. Dit moet een historisch onderzoeker altijd doen, maar in andere professionele digitale archieven is veel meer expertise en helderheid aanwezig. Er is geen sprake van een cultureel geheugen binnen The Internet Archive.

Daarnaast mist er veel expertise en reflectie op de bronnen, waardoor de bruikbaarheid niet gegarandeerd kan worden. Het is vooral een onderzoekwaardig archief, om deels net als dit onderzoek, te kijken naar hoe gebruikers het verleden willen laten zien en welke dingen zij belangrijk vinden om te bewaren. Dit kan verschillen van het hoogstpersoonlijke of het ideologisch sterk gekleurde. The Internet Archive is meer een gigantische opslagplaats of loods gevuld met willekeurige digitale producten door verschillende processen en personen, dan een zorgvuldig bekeken archief waar elk archiefstuk een bepaalde functie heeft. Dit maakt het lastig om daar gedegen onderzoek naar te doen.

Het doen van historisch onderzoek kan worden verbeterd door vooral veel meer de dialoog aan te gaan. Waar veel andere digitale archieven de gebruiker hebben betrokken bij het

professionele archief is het bij The Internet Archive aan te raden om vooral de professional te betrekken bij het gebruikersarchief. Hierdoor worden bronnen een stuk betrouwbaarder en kan er gewerkt worden aan een breed internationaal cultureel geheugen. Met enkele technische oplossingen en richtlijnen kan de selectie en ordening stukken verbeterd worden en helpen om mensen meer kennis te geven.

Hoofdstuk 7 Conclusie

Is het Nederlandstalige gedeelte van The Internet Archive nuttig voor het doen van historisch onderzoek? Om een antwoord te vinden op deze vraag is het belangrijk om te weten welke eisen er aan een digitaal archief worden gesteld. De archieftheorie spreekt voornamelijk over ordening en selectie binnen een archief. Ordening is nodig om het archief overzichtelijk in inhoud te houden. Selectie van archiefstukken door de archivariissen is nodig om in het fysieke archief ruimtegebrek tegen te gaan, maar ook om na te denken over het beeld van het archief. Dit laatste is onder postmoderne invloeden uitgebreid met de notie dat een archief een cultureel geheugen moet scheppen met zijn archiefstukken. Het archief bewaart een gedeelte van de samenleving en moet openstaan voor die samenleving. Selectie moet gebeuren in samenspraak met de gebruikers van het archief. Het archief is geen fysiek instituut meer, maar een stem binnen de maatschappij.

Dit verandert de rol van de professional. Deze moet veel kritischer het archief beschouwen. Zowel bij de selectie van nieuwe stukken, waarbij moet worden nagedacht hoe belangrijk deze zijn voor het cultureel geheugen alsmede de reflectie op oude stukken. Hebben zij nog een rol? Hoe worden ze gebruikt? Zeker in het digitale tijdperk waar veel gekopieerd en gedeeld kan worden, kunnen archiefstukken snel een eigen leven gaan leiden.

Digitale archieven bieden voor deze postmoderne visie een goede uitkomst. Door de digitalisering is de bereikbaarheid van archieven beter geworden. Gebruikers kunnen vanaf hun computer zichzelf toegang verschaffen tot de (openbare) archiefstukken. Daarnaast gebruiken gebruikers veel meer dan één archief en pakken archiefstukken uit vele digitale archiefopties. Dit sterkt de digitale archieven als databases met cultureel materiaal en geeft ze een veel bereikbaardere plek in de samenleving. De meest recente ontwikkeling in Nederland is het samen laten werken van professionele digitale archieven, archiefstukken van archief A verbonden met vergelijkbare archiefstukken uit archief B. Deze trend valt samen met de Web 2.0 en Web 3.0-invloeden binnen de digitale wereld.

De gebruikers spelen hierbij de belangrijkste rol. Zij kunnen in de professionele digitale archieven helpen bij het beschrijven van bronnen, zodat het voor de andere gebruikers vindbaar wordt. Daarnaast kunnen ze de inhoud van het archief en daarmee het culturele geheugen (digitaal) promoten. Dit kan onder andere door het hergebruiken van archiefstukken

in nieuwe digitale creaties, dit doormiddel van Creative Commons. Deze Commons helpen om de rechten van archiefstukken aan te geven.

Voor persoonlijke digitale archieven geldt hetzelfde als voor professionele openbare digitale archieven, ook hier zijn ordening, selectie en herinnering van belang. De digitale opslagmogelijkheden maken dat de hoeveelheid digitale persoonlijke documenten en berichten zowel online als offline heel groot kan zijn. Echter niet al deze berichten hebben voor de persoon die ze geschreven heeft een grote waarde. Daarnaast kunnen sommige berichten en foto's nadelig zijn voor de toekomst, terwijl ze toch altijd online openbaar bewaard blijven.

Ook voor persoonlijke digitale archieven geldt dat het handig is dat de selectie door de persoon zelf plaatsvindt. Dit voorkomt dat nabestaanden tegenover een groot of confronterend archief komen te staan. Dit stelt de gebruiker in staat om zijn eigen herinnering op te stellen en aan te laten geven welke digitale documenten voor hem van belang zijn of waren. Het actief beheren kan voorkomen dat beschamende digitale documenten online een eigen weg gaan.

The Internet Archive is één van de eerste digitale archieven, die de mogelijkheden aanbiedt om een persoonlijk archief te maken en beheren. Echter is het niet hun doel om persoonlijke archieven te verzamelen. In de eerste plaats is hun doel om alle (versies van) websites te bewaren om later te kunnen bekijken en onderzoeken. In de tweede plaats kwamen daar alle gedigitaliseerde en *digital born*-boeken, -audio en -video bij. Dit allemaal vanuit de doelstelling om alle kennis beschikbaar te maken voor iedereen. Er is nu een gigantisch archief verzameld van miljarden websites en miljoenen boeken, audio en video. Deze archiefstukken worden deels verzameld door bots, maar ook door gebruikersaccounts.

De verzamelde websites worden regelmatig teruggehaald via The WaybackMachine. Wetenschappelijk onderzoek naar The Internet Archive heeft zich vooral op die functie gefocust. Dit dan voornamelijk als tool voor onderzoek, verwijderde links in wetenschappelijke artikelen konden zo worden teruggehaald. De andere literatuur gaat vooral over de technische aspecten rondom de websites, het verzamelen daarvan en dergelijke. Er is bijna geen onderzoek gedaan naar de inhoud van The Internet Archive buiten de websites om. Het enige onderzoek, wat dit onderzoek heeft kunnen vinden, van Anya Boyle noemt een specifiek archief slecht geordend en erg persoonlijk.

Ondanks de doelstelling van The Internet Archive om niet per definitie persoonlijke archieven te verzamelen, tonen veel van deze gebruikersaccounts wel dit type bron. Dit werd duidelijk door het zoeken naar Nederlandse zoektermen in het grote archief. Dit verschilt van persoonlijke foto's tot promotiemateriaal voor hun creatieve uitingen. Daarnaast kan een actief gebruikersaccount erg zijn invloed laten zien op deze zoektermen. Voor de Nederlandse termen was dit vooral door een account wat actief was voor een historisch genootschap in Zandvoort en door een account wat fascistische literatuur wilt promoten.

Zijn deze bronnen in The Internet Archive het waard om historisch onderzoek naar te doen? Alhoewel er inhoudelijk zeker interessante bronnen bij zitten ontbreken in de basis toch de drie steekwoorden die binnen de archieftheorie zo belangrijk zijn: ordening, selectie en herinnering. Er is geen selectie, elke bron wordt in principe toegelaten. De ordening is (te) ruim in The Internet Archive, waardoor archieven te groot en divers worden. Door het ontbreken van deze twee is er geen ruimte om een lijn te kunnen trekken door de verschillende archiefstukken. Hiermee ontbreekt een cultureel geheugen binnen het archief. Door de internationaliteit en grootte van het archief komt een onderzoeker ook veel bronnen tegen die niet nuttig zijn.

Hier komt bij dat veel archiefstukken geen context en achterliggende expertise hebben. Soms zijn het stukken die uit een ander archief komen en niet meer up-to-date zijn. Waarbij het wenselijker is om contact op te nemen met het originele archief, vooral omdat die up-to-date is, meer expertise heeft en meer vergelijkbare stukken. De stukken die gebruikers hebben geüpload zijn vaak slecht beschreven. Onduidelijke titels, geen tot weinig context en soms zelfs foutieve informatie maken de controleerbaarheid van bronnen moeilijk binnen The Internet Archive. The Internet Archive is op deze manier vooral een grote verzameling van veel bronnen, maar heeft niet de expertise en kennis over deze bronnen. Bronnen met een sterk ideologisch karakter worden niet genuanceerd door andere bronnen of achtergrond. Dit maakt The Internet Archive ongeschikt voor het doen van historisch onderzoek, alhoewel een interessante bron over hoe gebruikers het verleden bewaren.

7.1 Suggestie verder onderzoek en reflectie

Dit onderzoek heeft gekeken naar de Nederlandse stukken binnen The Internet Archive en dit gekoppeld aan noties over professionele digitale archieven en daar geprobeerd om Nederland voorop te stellen. Dit onderzoek toepassen in andere landen of taalgebieden is een mogelijke

vervolgstap. Het Nederlandse gedeelte van The Internet Archive is niet geschikt voor historisch onderzoek, dit kan anders zijn als er gefocust wordt op bronnen uit andere landen of andere zoektermen. De focus lag in dit onderzoek op brede historische onderwerpen, de grote hoeveelheid bronnen van The Internet Archive lenen zich ervoor om binnen andere vakgebieden het wetenschappelijk nut te bekijken.

Voor zover door dit onderzoek gevonden is The Internet Archive het enige digitale archief wat een grote nauwelijks gecontroleerde rol aan de gebruiker geeft om openbaar stukken te mogen publiceren. De literatuur bespreekt altijd de gebruiker enerzijds in een gecontroleerde rol binnen het professionele digitale archief en anderzijds als beheerder/schepper van zijn persoonlijke digitale archief. In de literatuur komen deze twee verschillende soorten archieven weinig bij elkaar. Dit onderzoek is een eerste opzet om een archief, The Internet Archive, die dat wel doet te onderzoeken. Het is wenselijk om (internationaal) meer van dit soort archieven te vinden en te zorgen dat er meer verbindend wordt gekeken naar de professionele en persoonlijke digitale archieven die online te vinden zijn. Mocht er ooit een vergelijkbaar digitaal archief als The Internet Archive verschijnen dan is een vervolgonderzoek op dat archief zeker van belang.

Er is getracht om in dit onderzoek de verschillende opvattingen rondom digitale bronnen en archieven zo breed mogelijk te bespreken. Een tweetal begrippen/ontwikkelingen zijn hierbij benoemd, maar niet uitgebreid beschreven in de context van The Internet Archive. De eerste is digitaal erfgoed. Zeker het *digital born*-erfgoed in de vorm van websites is natuurlijk de kern van het oprichten van The Internet Archive. In de literatuur wordt deze term veel gebruikt en kan het verschillende dingen betekenen. De boeken van Google Books uit de negentiende eeuw vallen volgens de meeste definities onder de koepelterm digitaal erfgoed, alleen omdat ze in gedigitaliseerde vorm te lezen zijn. Een begrijpelijke definitie, maar nogal breed, vooral met steeds meer *digital born*-erfgoed in het vooruitzicht. Hoe grotere sociale online netwerken en berichtendiensten een plek gaan vinden in digitale archieven is weinig onderwerp van debat. Opmerkelijk want deze netwerken combineren de combinatie tussen persoonlijke en algemene bronnen, in eenzelfde, zelfs grotere hoeveelheid als The Internet Archive. Deze netwerken en hun functies zijn misschien nog wel meer het digitale erfgoed dan de losse gedigitaliseerde en *digital born*-stukken waar het wetenschappelijke debat nu vooral op focust.

De tweede ontwikkeling waar dit onderzoek eventueel te weinig aandacht aan heeft geschonken zijn de voornamelijk juridische ontwikkelingen rondom The Right to be

Forgotten. Dit is een zeer actueel onderwerp. Het in hoofdstuk 3 aangehaalde NRC-artikel over het vergeten worden in Google-zoekresultaten is verschenen tijdens het schrijven van dit onderzoek. Veel van de juridische besluiten die nu genomen worden omtrent persoonlijke data op het internet gaan van invloed zijn op het bewaar- en selectiebeleid van digitale archieven. Dit onderzoek heeft hier niet op gefocust, er van uitgaande dat alles wat op The Internet Archive staat, in ieder geval al lange tijd, er legaal op staat. Privacy en juridische uitspraken zijn onderwerpen die zeker voor de toekomst van het digitale archief actief gevolgd moeten worden.

Digitale archieven hebben nog veel uitdagingen voor de boeg. De digitale opslagmogelijkheden en de digitale productie buiten het archief blijven beiden groeien. Dit betekent veel en meer gedocumenteerde digitale bronnen voor de toekomst. Voor onderzoek naar deze bronnen blijft een kritische reflectie op de archiefstukken noodzakelijk. Ordening en selectie blijven van belang, als mede een dialoog tussen gebruiker en beheerder van het archief. Hierdoor kan onderzoek naar digitale archieven relevant en nuttig worden. Zonder ordening en selectie vervalt een digitaal archief in een massale verzamelplaats, waarbij context en controleerbaarheid van archiefstukken verdwijnen. Dit onderzoek heeft met zijn casus hiervan een voorbeeld willen geven, om deze dialoog relevant te maken.

Literatuurlijst

Boeken en artikelen

- Adriaans, W., J van den Berg, L. Breure en A. Melief, *Alles uit de kast. Op weg naar een nationaal investeringsprogramma digitale infrastructuur cultureel erfgoed.* (z.p. 1998).
- AlNoamany, Yasmin, Ahmed Alsum, Michele C. Weigle en Michael L. Nelson 'Who and what links to the Internet Archive' *Int J Digit Libr* (2014) 14, 101 – 115.
- Banks, Richard, 'Our Technology Heritage' in: Donald T. Hawkins (red.), *Personal Archiving: Preserving our Digital Heritage* (Medford 2013) 215 – 242.
- Barwick, Joanna, James Dearnley en Adrienne Muir, 'Playing games with Cultural Heritage: A comparative case study analysis of the current status of digital game preservation', *Games and Culture* 6 (2011) 4, 373 – 390.
- Benjamin, Walter en Harry Zohn (vert.) 'The Work of Art in the Age of Mechanical Reproduction', in: Hannah Arendt, *Illuminations* (z.p. 1969) 1-26.
- Blais, Gabrielle en David Enns 'From Paper Archives to People Archives: Public Programming in the Management of Archives.' *Archivaria* 31 (1990) 101-113.
- Boertjes, Erik, Almarima Jamakovic en Stephan Raaijmakers, 'Analyse van sociale media', *Stator* 3-4 (2011), 44-47.
- Bowman, Shayne en Chris Willis, *We Media, how audiences are shaping the future of news and information* (z.p. 2003).
- Boyle, Anya, *Archiving the Internet, what does it mean in practice?* (Leiden 2016).
- Bruns, Axel 'Prodisage: Towards a Broader Framework for User-Led Content Creations' *Proceedings Creativity & Cognition* 6 (2007) 1-7.
- Cook, Terry en Joan M. Schwartz, 'Archives, Records and Power: From (Postmodern) Theory to (Archival) Performance' *Archival Science* 2 (2002) 171 – 185.
- Debuysere, Stoffel, 'Onvoltooid Verleden Tijd, Brokstukken als Bouwstenen: Voortbouwen op de ruïnes van het audiovisuele geheugen.' in: Stoffel Debuysere e.a. (red.), *Bewaring en Ontsluiting van Multimediale Data in Vlaanderen: perspectieven op Audiovisueel erfgoed in het digitale tijdperk* (Leuven 2010) 125-139.
- Derrida, Jacques en Eric Prenowitz (vert.), 'Archive fever: A Freudian Impression' *Diacritics* 25 (1995) 2, 9-63.
- Dulong de Rosnay, Melanie en Andres Guadamuz 'Memory Hole or Right to Delist? Implications of the Right to be Forgotten for Web Archiving' *Reset Online* 6 (2017) 1-22.
- Edwards, Eli, 'Ephemeral to Enduring: The Internet Archive and Its Role in Preserving Digital Media' *Information Technology and Libraries* (2004) 3-8.
- Gehl, Robert, 'YouTube as archive, who will curate this digital wunderkammer?' *International journal of cultural studies* 12 (2009) 1, 43-60.
- Gibbs, Fred en Trevor Owens, 'The Hermeneutics of Data and Historical Writing', in: Jack Dougherty en Kristen Nawrotzki (red.), *Writing history in the digital age*, (Ann Arbor 2013) 159 – 170.
- Hollander, Frank den, 'Mulerius: erfgoed per tweet: Hoe een robot de digitale collecties doorzoekt...' *Pictogram* 2 (2016) 4, 6-7.
- Jaffe, Elliot en Scott Kirkpatrick 'The Architecture of the Internet Archive' *SYSTOR '09 Proceedings of SYSTOR 2009: The Israeli Experimental Systems Conference* artikel 11 (2009) 1-10.

- Jeurgens K.J.P.F.M., A.C.V.M. Bongenaar en M.C. Windhorst (red.), *Gewaardeerd verleden: bouwstenen voor een nieuwe waarderingsmethodiek voor archieven* (z.p. 2007).
- Jong, Annemieke de, *Het Boek van Zand, Naar een ontologie van het digitale archief* (Hilversum 2006).
- Keen, Andrew *The Cult of the Amateur: How blogs, MySpace, YouTube and the rest of today's user-generated media are killing our culture and economy* (Londen 2007).
- Leeuw, Sonja de, 'Het archief als netwerk, perspectieven op de studie van online televisie-erfgoed' *Tijdschrift voor Mediageschiedenis* 14 (2012) 1, 10-28.
- Lonkhuizen, Liza van 'Vergeetrecht' wordt op het web lang niet altijd toegekend', *NRC Handelsblad*, 11 mei 2017.
- Lusenet, Yola de *Preservation of digital heritage, draft discussion paper prepared for UNESCO* (z.p. 2002).
- Lusenet, Yola de, 'Tending the Garden or Harvesting the Fields: Digital Preservation and the UNESCO Charter on the Preservation of the Digital Heritage', *Library Trends* 56 (2007) 1, 164-182.
- Manoff, Marlene, 'Theories of the Archive from Across the Disciplines', *Libraries and the Academy* 4 (2004) 1, 9-25.
- Mayer- Schönberger, Viktor, *delete, The Virtue of Forgetting in the Digital Age* (Princeton 2009).
- Mostert, Kelly, 'Turning Historical Newsreels into Awesome Gif Animations' (versie 4 juli 2016), <https://www.beeldengeluid.nl/en/blogs/research-and-development/201607/turning-historical-newsreels-awesome-gif-animations> (25 mei 2017).
- Nesmith, Tom, 'Seeing Archives: Postmodernism and the Changing Intellectual Place of Archives' *The American Archivist* 65 (2002) 24 – 41.
- Netwerk Digitaal Erfgoed, *Nationale strategie digitaal erfgoed* (z.p. 2015).
- Nil, Bart de, 'Eigentijds beschrijven: Het contextualiseren van digitaal audiovisueel erfgoed in archieven' in: Stoffel Debuysere e.a. (red.), *Bewaring en Ontsluiting van Multimediale Data in Vlaanderen: perspectieven op Audiovisueel erfgoed in het digitale tijdperk* (Leuven 2010) 36 – 45.
- Nil, Bart de en Jeroen Walterus, 'Cultureel erfgoed en web 2.0', in: Bart de Nil en Jeroen Walterus (red.), *Erfgoed 2.0, nieuwe perspectieven voor digitaal erfgoed* (Brussel 2009) 9-17.
- Nulens, Gert, Eva van Passel en Joke Beyl 'Mag het wat actiever? De gebruiker als prosumer van een multimedia-archief', in: Stoffel Debuysere e.a. (red.), *Bewaring en Ontsluiting van Multimediale Data in Vlaanderen: perspectieven op Audiovisueel erfgoed in het digitale tijdperk* (Leuven 2010) 77 – 89.
- Nulens, Gert 'Virtuele Participatie', in: Bart de Nil en Jeroen Walterus (red.), *Erfgoed 2.0, nieuwe perspectieven voor digitaal erfgoed* (Brussel 2009) 109 – 120.
- Owens, Trevor 'When did we become users?' (versie 5 augustus 2011), <http://www.trevorowens.org/2011/08/when-did-we-become-users/> (25 mei 2017).
- Pietrobruno, Sheenagh, 'YouTube and the social archiving of intangible heritage' *New Media & Society* 15 (2013) 8, 1259 – 1276.
- Pitsillides S., J. Jefferies, en M. Conreen, 'Museum of the self and digital death: an emerging curatorial dilemma for digital heritage' in: Elisa Giaccardi (red.), *Heritage and social media. Understanding heritage in a participatory culture* (New York 2012) 56 – 68.
- Ridener, John, *From polder to postmodernism: an intellectual history of archival theory* (San Jose 2007).

- Rossaak, Eivind 'Analyzing archives in motion' *Journal of Scandinavian Cinema* 1 (2010) 1, 79 – 81.
- Rosenzweig, Roy, *Clio-wired: The future of the past in the digital age* (New York 2011).
- Schwartz, Joan M. en Terry Cook 'Archives, Records, and Power: The Making of Modern Memory' *Archival Science* 2 (2002) 1 – 19.
- Stobbe, J., C. de Maijer en L. Cavero 'Van small data naar big audience: Collecties Zuid-Afrikahuis in bewerking (2)' *Maandblad Zuid-Afrika* 92 (2015) 5, 110-111.
- Terranova, Terranova 'Free Labor: Producing Culture for the Digital Economy' *Social Text* 63 (2000) 18:2, 33-58.
- Thelwall, Mike en Liwen Vaughan, 'A fair history of the Web? Examining country balance in the Internet Archive' *Library & Information Science Research* 26 (2004) 162 – 176.
- Valois-Nadeau, Fanny, 'Rethinking heritagization through the digitization of familial archives' *Reset* 6 (2017) 1-16.
- Wilson, Lizette, 'One on one with Brewster Kahle, co-founder of the Internet Archive' (versie 21 oktober 2001), <http://www.bizjournals.com/sanfrancisco/stories/2001/10/22/newscolumn9.html> (19 juni 2017).
- Wolff, Robert S. 'The Historian's Craft, Popular Memory and Wikipedia', in: Jack Dougherty en Kristen Nawrotzki (red.), *Writing History in the Digital Age* (Ann Arbor 2012) 64-74.
- Wubs, Henrieke en Frank Huysmans, *Klik naar het verleden: een onderzoek naar gebruikers van digitaal erfgoed: hun profielen en zoekstrategieën* (Den Haag 2006).
- Ximm, Aaron 'Active Personal Archiving and The Internet Archive', in: Donald T. Hawkins (red.), *Personal Archiving: Preserving our Digital Heritage* (Medford 2013) 187 – 214.
- 'vroegste Nederlandse videogames in archief van Beeld en Geluid' (versie 12 februari 2016) <http://www.beeldengeluid.nl/nieuws/201602/vroegste-nederlandse-videogames-archief-van-beeld-en-geluid> (19 juni 2017).

Websites en bronnen The Internet Archive

Algemene websites en weblinks ter controle van bronnen The Internet Archive.

- Amsterdammuseum, project digitale stad <https://hart.amsterdam/nl/page/521/re-dds> (20 juni 2017).
- Creative Commons. 'The State of the Commons' <https://stateof.creativecommons.org/> (25 mei 2017).
- Bok, 'Nederlandstalige Bibliotheek Officieel Geopend', (versie 27 mei 2015) <https://web.archive.org/web/20170106121821/http://frissekijk.info/nederlandstalige-bibliotheek-officieel-geopend/> (versie WaybackMachine 6 januari 2017) (6 juni 2017).
 - o Originele link: <http://frissekijk.info/nederlandstalige-bibliotheek-officieel-geopend/> (niet werkend 6 juni 2017).
- Website EenVandaag <http://www.eenvandaag.nl/> (6 juni 2017).
- Genootschap Oud Zandvoort, <https://www.oudzandvoort.nl/> (6 juni 2017).

- Genootschap Oud Zandvoort, De Zandvoortsche Courant 2014
<https://www.oudzandvoort.nl/geschiedenis/digitale-kranten/572-de-zandvoortse-courant-2014> (6 juni 2017).
- The Internet Archive, <https://archive.org/> (16 juni 2017).
 - About The Internet Archive <https://archive.org/about/> (20 juni 2017).
 - The Live Music Archive, <https://archive.org/details/etree> (20 juni 2017).
 - Waybackmachine <https://archive.org/web/> (20 juni 2017).
- Openbeelden-project Beeld en Geluid, <https://www.openbeelden.nl/> (25 mei 2017).
 - Video Internationaal Torenspringen,
https://www.openbeelden.nl/media/2053/Internationaal_torenspringen (25 mei 2017).
- Website Socialistische Partij (SP), <https://www.sp.nl> (20 juni 2017).
 - SP, ‘Emile Roemer: ‘Er is genoeg voor iedereen’’
<https://www.sp.nl/video/emile-roemer-er-is-genoege-voor-iedereen> (6 juni 2017).
 - Foutieve link uit The Internet Archive-bron:
https://www.sp.nl/studio/1077003/thema::geld/20100209-vriendjespolitiek_van_het_old_boys_network.html (foutmelding 6 juni 2017).
 - Link waar The Internet Archive-bron waarschijnlijk naar refereert: SP, ‘Bedrijfsleven lobbyt tegen beperking aantal nevenfuncties door SP’ (versie 6 februari 2010) <https://www.sp.nl/nieuws/2010/02/bedrijfsleven-lobbyt-tegen-beperking-aantal-nevenfuncties-door-sp> (6 juni 2017).
- YouTube Statistics, <https://www.youtube.com/yt/press/statistics.html> (20 juni 2017).

Bronnen The Internet Archive

- 0619 Historische Tatsachen Nr. 53 Udo Walendy Enstellte Geschichte (1992, 44 S. Verboten)
<https://archive.org/details/0619HistorischeTatsachenNr.53UdoWalendyEntstellteGeschichte199244S.Verboten> (6 juni 2017).
- De Avonturen van Ome Keesie [VARA 26 febr 1941]
<https://archive.org/details/DeAvonturenVanOmeKeesievara26Febr1941> (6 juni 2017).
- De Brusselsche anti-slavernij conferentie 1889-1890,
<https://archive.org/details/debrusselschean00vissgoog> (25 mei 2017).
- Collectie Filmcollectief – Familie Koper in Nederlands-Indië
<https://archive.org/details/familie-koper-uit-nederlands-indie> (6 juni 2017).
- filmcollectief-SR-0067, <https://archive.org/details/filmcollectief-SR-0067> (6 juni 2017).
- filmcollectief-SR-0072 <https://archive.org/details/filmcollectief-SR-0072> (6 juni 2017).
- GOZ-PDF-H02181 <https://archive.org/details/GOZ-PDF-H02181> (6 juni 2017).
- Inside Boulder News – Flood Update
https://archive.org/details/Inside_Boulder_News_-_Flood_Update (6 juni 2017).
- Internationaal torenspringen, <https://archive.org/details/oi2053> (25 mei 2017).
- Jaarrapport Integratie 2013 <https://archive.org/details/JaarrapportIntegratie2013> (6 juni 2017).
- Jorge Valdes, Radio Nederland Wereldomroep
<https://archive.org/details/RNWMixEdicionEspecialJorgeValdes> (6 juni 2017).

- Keek op de Week – 96 = 3van7 – Centrum Democraten <https://archive.org/details/KeekOpDeWeek963van7CentrumDemocraten> (6 juni 2017).
- Konnijntje Official <https://archive.org/details/KonnijntjeOfficial> (6 juni 2017).
- mooie, <https://archive.org/details/mooie> (6 juni 2017).
- Nederlandse Eetbare Planten en Paddenstoelen Database <https://archive.org/details/NederlandseEetbarePlantenEnPaddenstoelenDatabase> (6 juni 2017).
- RTV Drenthe 2016-11-04 14:49, https://archive.org/details/archiveteam_videobot_RTV_Drenthe_20161104_1449 (25 mei 2017).
- RTV Drenthe 2017-05-23 09:19, https://archive.org/details/archiveteam_videobot_RTV_Drenthe_20170523_0919 (6 juni 2017).
- De slavernij in Suriname; of, Dezelfde gruwelen der Slavernij: die in de “Negerhut” geschetst ... <https://archive.org/details/deslavernijinsu00wolbgoog> (6 juni 2017).
- vriendjespolitiek <https://archive.org/details/vriendjespolitiek> (6 juni 2017).
- Wilders is Vele Malen Erger dan Janmaat <https://archive.org/details/WildersIsVeleMalenErgerDanJanmaat> (6 juni 2017).
- Wolff [gegraveerd door Floris Baltasarz] <https://archive.org/details/ned-kbn-all-00014740-001> (25 mei 2017).

Extra bronnen The Internet Archive met onduidelijke/ontbrekende/gekleurde metadata

- Aard van het Beest, De, <https://archive.org/details/AardVanHetBeestDeWilliamLutherPierce> (afkomstig uit Bitterburg archief) (20 juni 2017).
- De Beste Video's van DJ Jaap. <https://archive.org/details/DeBesteVideosVanDjJaap> (20 juni 2017).
- Cops In Holland Beat Up Jakes Protesting Against ‘Black Face’ (zwarte Piet), <https://archive.org/details/CopsInHollandBeatUpJakesProtestingAgainstBlackFacezwartePiet> (20 juni 2017).
- Masterscriptie.archive <https://archive.org/details/Masterscriptie.archive> (20 juni 2017).
- NL-HaNA, Kabinet Minister-President, 2.03.01, inv.nr. 3017 <https://archive.org/details/Nl-hanaKabinetMinister-president2.03.01Inv.nr.3017> (20 juni 2017).
- Santa and Black pedro in the Netherlands, https://archive.org/details/antonwithagen_gmail_Sint (20 juni 2017).
- Zwarte Piet, https://archive.org/details/ZwartePiet_98 (20 juni 2017).

Gebruikersaccounts The Internet Archive

- Arkiver, <https://archive.org/details/@arkiver> (6 juni 2017).
- B.Bitterburg https://archive.org/details/@b_bitterburg (6 juni 2017).
- Filmcollective Worldwide https://archive.org/details/@filmcollective_worldwide (6 juni 2017).
 - o Collectie Filmcollective Worldwide, https://archive.org/details/collectie_filmcollectief?&sort=-publicdate (6 juni 2017).

- Hans Steen, https://archive.org/details/@hans_steen (6 juni 2017).
- Joep Zander, <https://archive.org/details/@joepzander> (6 juni 2017).
- kimberlybmichalec <https://archive.org/details/@kimberlybmichalec> (6 juni 2017).
- magicshow, <https://archive.org/details/@magicshow> (6 juni 2017).
- Marzzz, <https://archive.org/details/@marzzz> (6 juni 2017).
- tinova, <https://archive.org/details/@tinova> (6 juni 2017).

Zoektermen The Internet Archive

- ‘Koloniale Oorlog’ <https://archive.org/search.php?query=Koloniale%20oorlog> (6 juni 2017).
- ‘Nederland’ <https://archive.org/search.php?query=Nederland> (6 juni 2017).
- ‘Nederlands-Indië’ <https://archive.org/search.php?query=Nederlands-Indi%C3%AB> (20 juni 2017).
- ‘Slavernij’ <https://archive.org/search.php?query=Slavernij> (6 juni 2017).
- ‘Tweede Wereldoorlog’
<https://archive.org/search.php?query=Tweede%20Wereldoorlog> (20 juni 2017).
- ‘Zwarte Piet’ <https://archive.org/search.php?query=Zwarte%20Piet> (20 juni 2017).

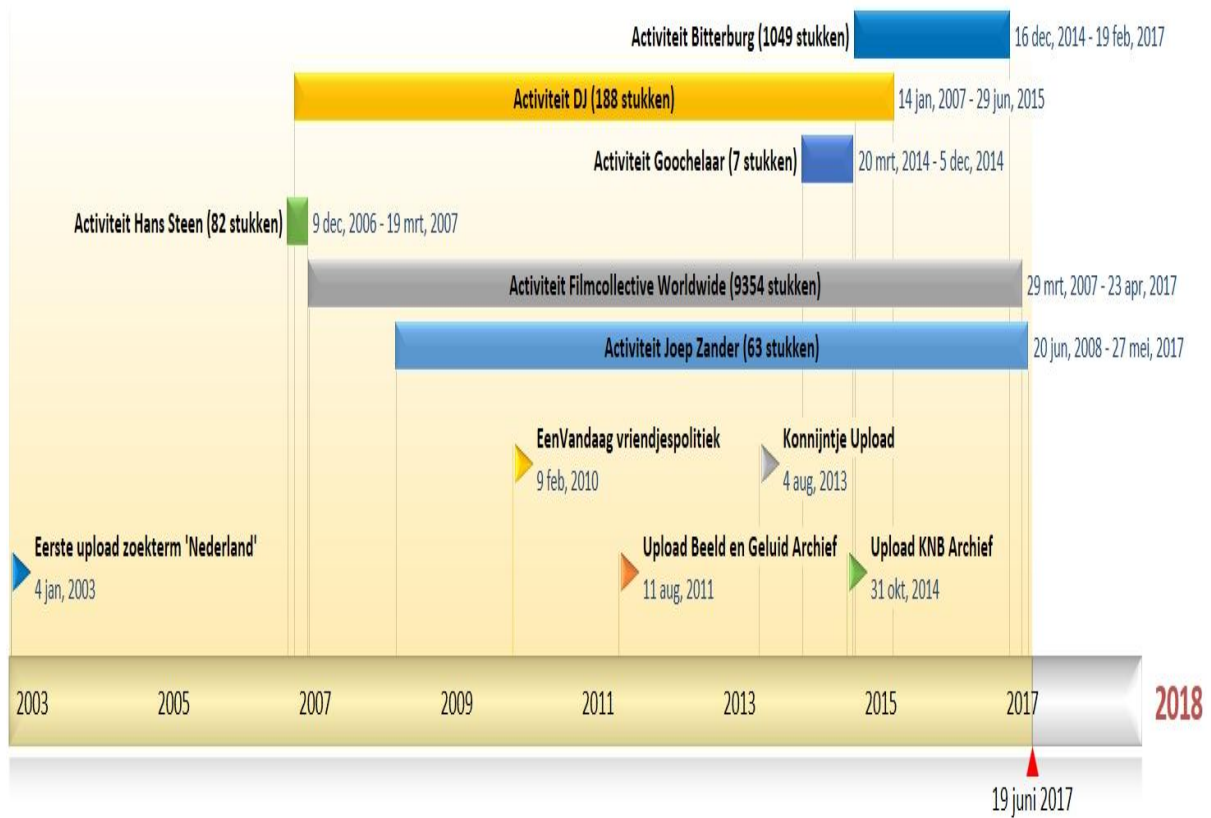
Alle afbeeldingen zijn screenshots gemaakt door de auteur.

Afbeelding voorblad:

https://commons.wikimedia.org/wiki/File:Internet_Archive_logo_and_wordmark.svg (23 juni 2017)

Bijlage 1 tijdslijnen activiteit gebruikersaccounts en zoektermen The Internet Archive

Tijdlijn 1 Activiteit Gebruikersaccounts en bronnen (4 januari 2003 – 19 juni 2017)



Tijlijn 2 Spreiding zoektermen The Internet Archive (4 januari 2003 – 19 juni 2017)

