

DIY Handboek voor Webarcheologie

Do It Yourself: Plan, graaf, reconstrueer en ontsluit!

Auteurs: Tjarda de Haan, Robert Jansma, Paul Vogel¹



Inleiding

Door de komst van het internet is onze samenleving fundamenteel veranderd. Het internet is met ons alledaagse leven verweven geraakt. Er is sprake van een sociale revolutie met een ongekende impact: werken, leren, spelen, communiceren en ons vermaken zijn onderhevig aan sociale en technologische innovaties en toepassingen. Deze ontwikkelingen gaan zo snel dat we ons permanent verhouden tot een veranderende werkelijkheid.

We verliezen de bronnen en informatie van de aanvang van ons digitale leven. De gemiddelde levensduur van een webpagina is 100 dagen. De eens zo populaire sites zoals Hyves, GeoCities en Altavista zijn alweer offline. Browsers en plug-ins raken verouderd waardoor webpagina's na verloop van tijd niet meer werken. Zo wordt het steeds lastiger om het verleden van het web te reconstrueren. Terwijl juist de beginjaren zo interessant zijn om te bestuderen. Er was, bij gebrek aan conventies, veel ruimte om te experimenteren met stijl en vorm. Google's "Chief Internet Evangelist" Vint Cerf waarschuwt zelfs voor een 'digital dark age': "Humanity's first steps into the digital world could be lost to future historians. We face a forgotten generation, or even a forgotten century. Our life, our memories (...) increasingly exist as bits of information - on our hard drives or in the cloud. But as technology moves on, they risk being lost in the wake of an accelerating digital revolution". De situatie is vergelijkbaar met hoe in het verleden bronnen van televisie, kranten en andere media verloren zijn gegaan. Kun je nu nog duidelijk maken hoe het internet eruit zag toen het ontstond? Wie de gebruikers waren? Hoe het gebruikt werd? Om ons digitale leven te kunnen bewaren voor de toekomst, moeten we nu ingrijpen zodat we geen cruciale informatie kwijtraken. De vergelijking met archeologie gaat op; oude systemen en interfaces moeten 'opgegraven' worden om het verleden te kunnen reconstrueren.

In 2017 is het 23 jaar geleden dat de eerste virtuele gemeenschap in Nederland, De Digitale Stad (DDS), werd gelanceerd. De Digitale Stad was voor veel van haar bewoners de eerste kennismaking met de mogelijkheden van het internet. Het is een onlosmakelijk onderdeel van de vroege internetcultuur van Nederland en speelt een belangrijke rol in de geschiedenis van Amsterdam, inmiddels het grootste internetknooppunt van de wereld. Binnen het project "De Digitale Stad Herleeft" hebben Waag Society, Amsterdam Museum, Nederlands Instituut voor Beeld en Geluid en de Universiteit van Amsterdam dit probleem geadresseerd. Dit consortium van erfgoedinstellingen, creatieve industrie en wetenschap heeft de afgelopen twee jaar belangrijke stappen gezet om De Digitale Stad voor huidige en toekomstige generaties te bewaren en ontsluiten.

Het meest bijzondere resultaat tot nu toe in het project is dat studenten van de UvA en VU, in samenwerking met voormalige DDS-medewerkers, erin zijn geslaagd twee reconstructies te (her)bouwen. Hiermee is de derde versie van DDS weer werkend gemaakt! Voor het ontsluiten van het erfgoed voor het publiek is een haalbaarheidsstudie gedaan die in de toekomst kan leiden tot een museale presentatie. Daarnaast heeft het project veel kennis opgeleverd over digitale preservatie en webarcheologie. Voor dit baanbrekende werk heeft het project in 2016 de prestigieuze Digital Preservation Award gewonnen als beste case study voor webarcheologie. DDS zal worden opgenomen in het e-depot van Beeld en Geluid. Hiermee is het behoud van dit erfgoed geborgd en is tevens ook de toegankelijkheid ervan voor wetenschappelijk onderzoek gegarandeerd. Als belangstellenden De Emulatie (een van de twee reconstructies) en de overige gevonden data willen bekijken, kunnen ze in de toekomst een verzoek indienen bij de stichting Beheer Digitaal Erfgoed DDS bestaande uit de huidige projectpartners, die waakt over het (her)gebruik van de DDS-data. De stichting beoordeelt de aanvragen van de onderzoekers.

Via 'FREEZE! Red en behoud ons digitaal erfgoed' wordt de opgedane kennis beschikbaar gesteld voor het erfgoedveld. 'FREEZE!' Bestaat uit drie onderdelen. Het 'Do It Yourself Handboek voor Webarcheologie' geeft op basis van de DDS casus praktische informatie over webarcheologie en is direct toepasbaar voor erfgoedprofessionals. Het 'Haalbaarheidsonderzoek Presentatie' beschrijft de resultaten van een studie naar de mogelijkheden om vroege Nederlandse internetcultuur te presenteren in een museale context. Het Manifest roept erfgoedinstellingen op om digitale preservatie en indien nodig webarcheologie ter hand te nemen zodat ons digitaal erfgoed niet verloren gaat.

Het consortium kijkt met gepaste trots terug op het project en hoopt dat 'FREEZE!' erfgoedinstellingen inspireert om na te denken over het veiligstellen van vroege internetcultuur. Speciale dank gaat uit naar de partijen die het onderzoek mogelijk gemaakt hebben: Mondriaan Fonds, Nationale Coalitie Digitale Duurzaamheid/Netwerk Digitaal Erfgoed, Prins Bernhard Cultuurfonds en Stimuleringsfonds Creatieve Industrie.

Colofon

Aan dit project hebben meegewerkt:

Projectgroep

Gerard Alberts (Universiteit van Amsterdam)
Judith van Gent (Amsterdam Museum)
Tjarda de Haan (Amsterdam Museum)
Robert Jansma (Universiteit van Amsterdam)
Lizzy Komen (Nederlands Instituut voor Beeld en Geluid)
Pam de Sterke (Waag Society)
Erwin Verbruggen (Nederlands Instituut voor Beeld en Geluid)
Robin van Westen (Waag Society)

Stuurgroep

Judikje Kiers (Amsterdam Museum)
Jan Müller (Nederlands Instituut voor Beeld en Geluid)
Julia Noordegraaf (Universiteit van Amsterdam)
Johan Oomen (Nederlands Instituut voor Beeld en Geluid)
Marijke Oosterbroek (Amsterdam Museum; vervanger)
Marleen Stikker (Waag Society)

En vele anderen, waaronder:

Gijs Boerwinkel (Waag Society)
Merel Boes (Waag Society)
Oud-DDS bewoners, oud-DDS werknemers en toeristen uit DDS.
Dick van Dijk (Waag Society)
Theun van den Doel (Amsterdam Museum)
Michaël van Eeden (oud-DDS-er)
Astrid Fiddelaers (Amsterdam Museum)

Joost Flint (oud-DDS-er)
Arnoud Goos (Nederlands Instituut voor Beeld en Geluid)
Jana Hardt (Nederlands Instituut voor Beeld en Geluid)
Laura van Hasselt (Amsterdam Museum)
Christine van den Horn (Waag Society)
Maartje Hülsenbeck (Nederlands Instituut voor Beeld en Geluid)
Mieke Lauwers (Nederlands Instituut voor Beeld en Geluid)
Robin van der Markt (Universiteit van Amsterdam)
Edo Mulder (Amsterdam Museum)
Marcel Ras (Nationale Coalitie Digitale Duurzaamheid/Netwerk Digitaal Erfgoed)
Josefien Schuurman (Nederlands Instituut voor Beeld en Geluid)
Marjolein Steeman (Nederlands Instituut voor Beeld en Geluid)
Studenten van de Universiteit van Amsterdam/Vrije Universiteit Amsterdam
Kees Teszelszky (Koninklijke Bibliotheek)
Bart Tunnissen (Waag Society)
Paul Vogel (oud-DDS-er)
Jesse de Vos (Nederlands Instituut voor Beeld en Geluid)
Carolien van Zuilekom (Nationale Coalitie Digitale Duurzaamheid/Netwerk Digitaal Erfgoed).

Een speciale dank gaat uit naar

Studenten informatica: Ronald Bethlehem, Thomas Koch, Millen Mortier, Kishan Nirghin, Tim Veenman, Marc Went.

<http://hart.amsterdam/freeze>

Powered by...



NATIONALE
COALITIE
DIGITALE
DUURZAAMHEID

NETWERK
DIGITAAL
ERFGOED



stimulerings
fonds
creatieve
industrie

Data verwelken en websites vergaan, maar
De Digitale Stad blijft eeuwig bestaan.

¹ Contact:

Tjarda de Haan: info@bitsandbytesunited.com (gast-conservator digitaal erfgoed, webregisseur),

Robert Jansma: R.Jansma.nl@gmail.com (student en assistent-docent),

Paul Vogel: spam@wateenlangedomeinnaamdiesvastnietech.nl (systeembeheerder, programmeur).

Inhoudsopgave

1	Introductie	7
2	Voor wie is dit handboek?	13
3	Hoe is dit handboek opgezet?	15
4	Getting started	19
5	Plannen	23
5.1	Beschrijf de doelstellingen	24
5.2	Doe bronnenonderzoek	26
5.3	Beschrijf het digitale object en het belang ervan	27
5.4	Breng in kaart wat je nodig hebt	33
5.5	Maak een plan van aanpak	40
5.6	Stel een begroting op	40
5.7	Zet een communicatieplatform op	41
6	Graven	45
6.1	Maak een inventarisatie van gewenste objecten	46
6.2	Maak je plan wereldkundig	46
6.3	Maak vondstkaarten	48
6.4	Breng mensen, verhalen en objecten bij elkaar	49
6.5	Garandeer de data-integriteit van je vondsten	60
6.6	Stel je vondsten veilig	61
7	Reconstrueren	65
7.1	Beschrijf de uitgangspunten	66
7.2	Onderzoek de ethische kwesties	67
7.3	Start met de voorbereidende werkzaamheden	70
7.4	Reconstrueer: emuleren of restaureren?	77
7.5	Duid de data	84

8	Ontsluiten	89
8.1	Archiveer: statisch of dynamisch?	90
8.2	Zoek een partner	91
8.3	Beschrijf de gewenste preserveringsscenario's	92
8.4	Geef toegang aan doelgroepen	95
8.5	Documenteer, documenteer, documenteer	95
9	Let the Bytes Free!	97
9.1	Toekomstig onderzoek	101
9.2	Deel jouw onderzoek	101
10	Bijlage I: Bronnen	103
11	Bijlage II: Woordenlijst	107
12	Bijlage III: Webarcheologische tools	111
13	Colofon	115

01

Introductie

“Humanity’s first steps into the digital world could be lost to future historians. We face a forgotten generation, or even a forgotten century”.²

Vint Cerf, 2015

Data vormen de nieuwe klei, scripts de nieuwe schoppen en het web is de allerjongste kleilaag die we gaan ontginnen. Webarcheologie is een nieuw gebied in e-cultuur waarbij we relatief nieuwe (born digital) objecten die nog niet zo lang geleden verloren lijken te zijn gegaan, opgraven en reconstrueren met nieuwe (digitale) gereedschappen. Zowel de materie als de methoden om ons digitale verleden op te graven en te reconstrueren zijn piepjong en nog volop in ontwikkeling. In dit handboek geven wij een eerste aanzet tot een beschrijving van webarcheologie. We gaan aan de slag, bit voor bit, byte voor byte. We laten niet alleen zien hoe we met digitale archeologie een digitale omgeving proberen op te graven uit het verleden, maar we laten ook zien hoe we onze archeologische vondsten veiligstellen en duurzaam proberen te bewaren voor toekomstige generaties.

Help! Ons digitaal erfgoed gaat verloren

Ons digitaal erfgoed dreigt verloren te gaan. Het behoud van onze digitale schatkamer en de blijvende toegang voor toekomstige generaties staan op het spel en vereisen onmiddellijke zorg en actie. Internetpionier van het allereerste uur Vint Cerf waarschuwde in 2015 al voor een 'digital dark age': “Our life, our memories (...) increasingly exist as bits of information - on our hard drives or in the cloud. But as technology moves on, they risk being lost in the wake of an accelerating digital revolution”.³ Bijna vijftien jaar geleden luidde UNESCO reeds de noodklok voor het mogelijke verlies en erkende de belangrijke historische waarde van ons (born) digitaal erfgoed als “unieke bronnen van de menselijke kennis en expressie”.⁴ Als we nu niet ingrijpen, raken we cruciale informatie over ons moderne leven kwijt. Hoewel de digitale (r)evolutie ons leven in de laatste decennia ingrijpend heeft veranderd, dreigt ons digitale verleden verloren te gaan. Of, erger, het is al weg.

² “Google boss warns of 'forgotten century' with email and photos at risk”, door Ian Sample, science editor The Guardian (13 februari 2015) <https://www.theguardian.com/technology/2015/feb/13/google-boss-warns-forgotten-century-email-photos-vint-cerf>.

³ “Google boss warns of 'forgotten century' with email and photos at risk”, door Ian Sample, science editor The Guardian (13 februari 2015) <https://www.theguardian.com/technology/2015/feb/13/google-boss-warns-forgotten-century-email-photos-vint-cerf>.

⁴ “Charter on the Preservation of Digital Heritage”, door UNESCO http://portal.unesco.org/en/ev.php-URL_ID=17721&URL_DO=DO_TOPIC&URL_SECTION=201.html

Als webarcheologen dalen wij af in de digitale erfgoedschatkamer van het internet. We gaan op zoek naar het digitale geheugen van het web. Het World Wide Web (het web) is 28 jaar geleden ontstaan⁵ en manifesteert zich in de vorm van websites,⁶ bestanden, hyperlinks en applicaties. Het web is in zeer korte tijd geavanceerder en complexer geworden.⁷ Websites zijn uitgegroeid van statische webobjecten tot dynamische webobjecten met (een veelheid aan) externe afhankelijkheden. Vergelijk de webobjecten met een zwarte doos: de werking van de zwarte doos is afhankelijk van het gedrag van het object, de gebruiker en van andere dynamische webobjecten. Een statisch webobject is voor iedereen hetzelfde, bijvoorbeeld een website opgebouwd uit kant-en-klare HTML-pagina's zoals een simpele bedrijfspagina. Een dynamisch object is afhankelijk van het gedrag van een gebruiker en diens interactie met de site, bijvoorbeeld bij een winkelmandje in een webshop of bij internetbankieren. Hierdoor is het behouden en het archiveren van het web steeds complexer geworden. De gemiddelde levensduur van een webpagina is 100 dagen.⁸ De vroegste resten van het web vinden we vaak alleen nog maar terug in webarchieven.

In de jaren negentig, de pionierstijd van het web, maakte een aantal instituten, waaronder het Internet Archive⁹ in de Verenigde Staten en de nationale bibliotheken van onder andere Groot-Brittannië en Australië,¹⁰ een begin met het archiveren van websites.¹¹ Het meest bekende webarchief, het Internet Archive, werd in 1996 in opgericht. Het Internet Archive wil universele toegang bieden tot álle menselijke kennis en bewaart inmiddels petabytes aan digitaal erfgoed in de vorm van boeken, films, software, muziek en websites. Bij webarchivering worden onderdelen van het web verzameld (ook wel 'web harvesten' of 'crawlen' genoemd) en opgeslagen in een webarchief formaat. Elementen uit deze webcollecties worden vervolgens weer toegankelijk gemaakt voor (her)gebruik.¹² Het Internet Archive presenteert de gearchiveerde websites (inmiddels miljarden) via de Wayback Machine in de vorm van 'snapshots', oftewel momentopnames van webpagina's. Deze webarchieven zijn ware goudmijnen voor onderzoekers.

⁵ <https://www.w3.org/webat25/about/history> en <http://home.cern/topics/birth-web>.

⁶ Niels Brügger onderscheidt vijf lagen in het webdomein: 1. elementen van een website (onder andere broncode, tekst, plaatjes en stylesheets), 2. afzonderlijke webpagina's (alle elementen die onder een zeker webadres te vinden zijn), 3. unieke website (alle webpagina's die onder een bepaalde domeinnaam te vinden zijn), 4. webomgeving (alle sites die met een bepaalde website verbonden zijn door middel van hyperlinks) en 5. wereldwijde web (alle websites die op een bepaald moment online zijn en bereikbaar zijn). Lees: The "Web as History: Using Web Archives to Understand the Past and the Present". UCL Press, 2017. Niels Brügger, Ralph Schroeder (editors). En "De oogst van de digitale Hollandse akker. Webarchivering in Nederland" door Kees Teszelszky in: Overheidsdocumentatie Online, juli/augustus 2017 <http://www.od-online.nl/artikel/od-juli-augustus-2017/de-oogst-van-de-digitale-hollandse-akker>.

⁷ Een website is een verzameling webpagina's met gegevens, zoals tekst, afbeeldingen of video's. De website staat op een of meer webserver. Webpagina's bestaan uit HTML-code en kun je bekijken in een webbrowser. HTML staat voor Hyper Text Markup Language. Door hypertext links in de HTML-code kun je van de ene pagina naar de andere navigeren. Gopher is de voorloper van het wereldwijde web, en werd gebruikt voor het zoeken en ophalen van documenten, ontworpen voor het internet. Toen Mosaic, de eerste webbrowser met een grafische gebruikersomgeving, verscheen in 1993 veranderde dit. De populariteit van de Gopher daalde en er ontstonden steeds meer websites. In het begin waren websites nog vrij eenvoudig; deze worden ook wel passieve of statische HTML pagina's genoemd. Een website toonde informatie (in opmaakvorm) en beelden. Later konden ook 'dynamische documenten' gemaakt worden. Vanaf 1995 gebeurde dit bijvoorbeeld met CGI, de common gateway interface. Met CGI kan via de webbrowser 'dynamisch' gegevens opgevraagd worden van de webserver. Hierdoor werd steeds meer mogelijk, bijvoorbeeld aanmelden op een website of het samenstellen van de inhoud van een website met gegevens uit een database. Deze websites worden ook wel actieve of dynamische pagina's genoemd.

Bij meer complexe, dynamische webobjecten bieden de huidige webarchieven geen uitkomst. Webarchieven bevatten 'snapshots' van statische websites. De dynamische webobjecten, die afhankelijk zijn van het gedrag van de gebruiker en het object, zijn niet eenvoudig te 'crawlen'. Neem bijvoorbeeld websites waar moet worden ingelogd. Een 'webcrawler' kan de pagina's na het inlogscherf niet zien. De webcrawler heeft geen account en kan niet inloggen. Daarnaast krijgt elke gebruiker 'eigen' unieke content te zien; webpagina's worden namelijk dynamisch gemaakt voor de gebruiker. Voor dynamische webobjecten heb je een andere aanpak nodig en moet je dieper graven.¹³

Als websites niet meer online zijn, dan moet op een andere manier onderzoek worden gedaan, die nog het meest lijkt op archeologisch onderzoek. Je moet dan eerst het webobject 'opgraven' en reconstrueren. Met deze aanpak ga je verder dan het maken van een momentopname zoals bij webarchivering: je gaat net als een archeoloog op zoek naar sporen die met het blote oog niet te zien zijn en niet door traditionele historische methoden kan worden gevonden. Een voordeel van deze methode is dat het publiek, na het opgraven en het reconstrueren, de oorspronkelijke werkende website weer kan bekijken en ervaren zoals het ooit online was. Net als bij archeologie moet je hiervoor 'de grond in' en graven. De overblijfselen, de gevonden scherven en andere artefacten, moet je eerst weer in elkaar lijmen, de sporen interpreteren en de oorspronkelijke digitale toestand van object, bits en bytes reconstrueren. Op deze manier verwerf je inzicht in de werking, het gebruik en de betekenis van het object en kun je het verleden reconstrueren en duiden. Het digitale equivalent van archeologie is dus webarcheologie.

⁸ "The Average Lifespan of a Webpage", door Mike Ashenfelder (8 november 2011):

<https://blogs.loc.gov/thesignal/2011/11/the-average-lifespan-of-a-webpage/> en Fixing Broken Links on the Internet, door Alexis Rossi (25 oktober 2013): <https://blog.archive.org/2013/10/25/fixing-broken-links/>

⁹ Internet Archive: <https://archive.org/> en de Wayback Machine: <http://web.archive.org/>. Bekijk ook: de FAQ op http://archive.org/about/faqs.php#The_Wayback_Machine en "The Internet Archive Documentary" door Jonathan Minard: <https://archive.org/details/youtube-1WxGDOeS-s4>.

¹⁰ <http://www.ncdd.nl/kennis-en-advies/expertgroepen/expertgroep-webarchivering/>.

¹¹ "The Web as History: Using Web Archives to Understand the Past and the Present". UCL Press, 2017. Niels Brügger, Ralph Schroeder (editors).

¹² Zie: <http://netpreserve.org/web-archiving/>. Het International Internet Preservation Consortium (IIPC) is in 2003 opgericht door twaalf nationale bibliotheken met als doel: 'preserving internet content for future generations'.

¹³ In Nederland is er een begin gemaakt met onderzoek naar het duurzaam behoud en beheer van born digital erfgoed voor de domeinen film, fotografie, architectuur en kunst. Het bewaren van nieuwe media en complexe born digital objecten zoals websites, games en apps, vielen echter buiten de scope. Lees: "Generieke workflows born digital erfgoed. Behoud van born digital erfgoed in Nederland: film, fotografie, architectuur, kunst". Onderzoek van de Culturele Coalitie Digitale Duurzaamheid naar born digital erfgoed, door Gaby Wijers en Hannah Bosma. (2015). http://ncdd.nl/site/wp-content/uploads/2015/10/CCDD-BornDigitalOnderzoek-def_2015.pdf.

Case study: 'De Digitale Stad Herleeft'

Om handen en voeten te geven aan het nieuwe vakgebied webarcheologie zetten we onze theorie in dit handboek meteen om in de praktijk. Dat doen we aan de hand van een case study: De Digitale Stad (DDS). Drieëntwintig jaar geleden verrees een unieke stad bestaande uit computers, modems en telefoonkabels. Op 15 januari 1994 opende De Digitale Stad in Amsterdam haar virtuele poorten. DDS, de eerste virtuele stad in de wereld, maakte het internet (gratis) toegankelijk voor het grote publiek in Nederland. Maar net als veel andere steden in de wereldgeschiedenis verdween ook deze stad. In 2001 werd De Digitale Stad offline gehaald en verging de stad als een virtueel Atlantis. Daarom vonden het Amsterdam Museum, Waag Society, Universiteit van Amsterdam en het Nederlands Instituut voor Beeld en Geluid het hoog tijd om DDS 'op te graven' en de digitale archeologische resten van DDS veilig te stellen voor de toekomst. Het project dat we hiervoor hebben opgezet heet 'De Digitale Stad Herleeft'.¹⁴ Het project fungeert niet alleen als een case study voor webarcheologie, het geeft tevens een eerste aanzet tot dynamisch webarchiveren, het duurzaam opslaan van en toegang bieden tot het (opgegraven en gereconstrueerde) werkende digitale object. In het 'DIY (Do It Yourself) Handboek voor Webarcheologie' delen we onze kennis en ervaringen die we opdoen tijdens het project en hopen hiermee de drempels te verlagen voor toekomstige webarcheologische projecten. We proberen antwoord te geven op de vragen: Hoe graaf je De Digitale Stad op en transformeer je het van een virtueel Atlantis naar een virtueel Pompeï? Hoe reconstrueer, bewaar, ontsluit en presenteer je DDS (born digital) data op een duurzame manier voor toekomstige generaties?

¹⁴ Dit 'DIY Handboek voor Webarcheologie' maakt deel uit van de eindpublicatie van het project 'DDS Herleeft': 'FREEZE! Red en behoud ons digitaal erfgoed. De Digitale Stad als case study voor webarcheologie'. In deze eindpublicatie zijn ook opgenomen: het 'Haalbaarheidsonderzoek Presentatie', waarin de resultaten van een studie naar de mogelijkheden om vroege Nederlandse internetcultuur te presenteren in een museale context, zijn beschreven; en het manifest 'FREEZE! A manifesto for safeguarding and preserving digital-born heritage', waarin erfgoedinstellingen worden opgeroepen om digitale preservering -en indien nodig webarcheologie- ter hand te nemen zodat ons digitaal erfgoed niet verloren gaat.

Powered by...

Dit handboek had niet geschreven kunnen worden zonder de enorme inspanningen en gulle bijdragen van oud-DDS bewoners, oud-DDS werknemers, toeristen uit DDS, vrijwillige webarcheologen, studenten van de Universiteit van Amsterdam/Vrije Universiteit Amsterdam en alle leden van ons projectteam bestaande uit het Amsterdam Museum, Waag Society, Universiteit van Amsterdam en het Nederlands Instituut voor Beeld en Geluid. Een speciale dank gaat uit naar studenten informatica: Marc Went, Thomas Koch, Ronald Bethlehem, Millen Mortier, Kishan Nirghin en Tim Veenman.

Het handboek is tot stand gekomen dankzij de financiële ondersteuning van het Prins Bernhard Cultuurfonds, Stimuleringsfonds Creatieve Industrie, de Nationale Coalitie Digitale Duurzaamheid (NCDD), het Netwerk Digitaal Erfgoed en het Mondriaan Fonds. Boven alles heeft de enthousiaste inzet en geweldige bijdragen van oud-DDS bewoners, oud-DDS werknemers en aangesloten webarcheologen, ons project ondersteund en ons gesterkt in onze missie. We bedanken de Digital Preservation Coalition en de NCDD voor het belonen van ons project met de internationale 'Digital Preservation Award 2016' in de categorie 'The National Archives Award for Safeguarding the Digital Legacy'.

02

Voor wie is dit handboek?

Dit handboek is bedoeld voor iedereen die aan de slag wil met webarcheologie en het archiveren van dynamische websites en dat direct in de praktijk wil brengen, van erfgoedprofessionals van collectiehoudende instellingen tot cultuurnerds met een missie. We laten zien hoe jij een webarcheologisch project kunt opzetten voor het duurzaam opgraven en ontsluiten van een verloren gewaand born digital object en hoe jij een relevant plan schrijft: *Wat wil je opgraven en behouden, waarom, door wie en voor wie en voor hoe lang.*

Je hebt geen technische kennis nodig om aan de slag te gaan. Wel bieden we naast de praktische handvaten voor erfgoedprofessionals en -liefhebbers, een technische verdieping voor technici die kunnen helpen bij de uitvoering.

03

Hoe is dit handboek opgezet?

Het 'DIY Handboek voor Webarcheologie' is een stap-voor-stap handleiding met praktische handvaten voor het opzetten en uitvoeren van een webarcheologisch project. Webarcheologie bestaat, in een notendop, uit een aantal processen: plannen, graven, reconstrueren en ontsluiten. Die beschrijven we per hoofdstuk. Onze case study 'De Digitale Stad Herleeft' loopt als rode draad door het handboek. In onze case study geven we antwoord op de vragen: Hoe graaf je De Digitale Stad op en transformeer je het van een virtueel Atlantis in een virtueel Pompeï? Hoe reconstrueer, bewaar, ontsluit en presenteer je DDS materiaal op een duurzame manier voor toekomstige generaties?

We nodigen alle (toekomstige) webarcheologen uit om zelf aan de slag te gaan en jullie kennis en ervaringen met ons te delen! Dig in en Do It Yourself! Plan, graaf, reconstrueer, bewaar, ontsluit, presenteer en deel jouw kennis!

Leeswijzer

In het volgende **hoofdstuk vier** 'Getting started' beschrijven we het belang van digitale objecten en leggen we uit wat digitaal erfgoed is en waarom born digital materiaal kwetsbaar is. Tot slot beschrijven we de aanleiding tot het project 'DDS Herleeft'.

In **hoofdstuk vijf tot en met acht** gaan we aan de slag. We voeren je mee naar het legendarische pand aan de Prins Hendrikkade 194 in Amsterdam, we maken je deelgenoot van de eerste webarcheologische opgravingen tijdens de 'Grave Diggers Party', we nemen je mee op onze ontdekkingstocht door diepe lagen rauwe data, naar bijzondere ontmoetingen met de oorspronkelijke bewoners uit De Digitale Stad, op zoektocht naar exotische tape readers en servers en op reis naar meer digitale avonturen. Per hoofdstuk diepen we de verschillende fasen van webarcheologie uit.

Hoofdstuk vijf 'Plannen': Schrijf een plan, beschrijf de doelstellingen, doe bronnenonderzoek, beschrijf het digitale object en het belang ervan, breng in kaart wat je nodig hebt, maak een plan van aanpak, stel een begroting op, zet een communicatieplatform op.

Hoofdstuk zes ‘Graven’: Maak een inventarisatie van gewenste objecten, maak je plan wereldkundig, maak vondstkaarten, breng mensen, verhalen en objecten bij elkaar, garandeer de data-integriteit van je vondsten, stel je vondsten veilig.

Hoofdstuk zeven ‘Reconstrueren’: Beschrijf de uitgangspunten, onderzoek de ethische kwesties, start met de voorbereidende werkzaamheden, reconstrueer: emuleren of restaureren?, duid de data.

Hoofdstuk acht ‘Ontsluiten’: Archiveer: statisch of dynamisch?, zoek een partner, beschrijf de gewenste preservingsscenario’s, geef toegang aan doelgroepen, documenteer, documenteer, documenteer.

In **hoofdstuk negen** ‘Let the Bytes Free!’ beschrijven we onze bevindingen en geven we een eerste aanzet voor toekomstig onderzoek.

In **bijlage I** ‘Bronnen’ geven we een overzicht met relevante publicaties en bronnen.

In **bijlage II** ‘Woordenlijst’ leggen we een (beperkt) aantal begrippen uit en verwijzen we je naar een aantal websites met een uitgebreide woordenlijst of leeromgeving.

In **bijlage III** ‘Webarcheologische tools’ geven we een overzicht van onze gereedschapskist.

De verschillende elementen per hoofdstuk:

- **Stappenplan:** Elk hoofdstuk vanaf ‘Getting started’ begint met een stappenplan. Dit is een checklist voor dat onderdeel van het proces. Je kunt deze volgen en -uiteeraard- aanvullen waar nodig voor jouw specifieke project.
- **Tips:** Tussendoor geven we praktische tips.
- **‘Onze aanpak: case study ‘DDS Herleeft’**’: Na de algemene webarcheologische uitleg delen we onze kennis en ervaringen uit het project ‘DDS Herleeft’. Deze stukjes herken je aan de kopjes ‘Onze aanpak: case study ‘DDS Herleeft’.
- Onder de kopjes **‘Technische verdieping’** belichten we de technische aspecten die vooral voor technici bedoeld zijn, maar die iedereen natuurlijk kan lezen!



04

Getting started

Objecten spelen voor musea en collectiehoudende instellingen een cruciale rol. ‘Dingen’ brengen het verleden tot leven en maken onze geschiedenis tastbaar. Inmiddels ontstaan er steeds meer objecten met een digitaal karakter, zoals websites, apps en games. Net als bijvoorbeeld film, fotografie, literatuur en muziek vormen deze digitale objecten een deel van ons erfgoed, ons dígitale erfgoed.

Wat verstaan we onder digitaal erfgoed? Digitaal erfgoed is de digitale verschijningsvorm van cultureel erfgoed en kent drie vormen:¹⁵

1. **Gedigitaliseerd erfgoed:** Erfgoed dat van origine niet digitaal is, maar waarvan met digitalisering een reproductie is gemaakt (bijvoorbeeld een scan van een boek).
2. **Digitale informatie over erfgoed:** Bijvoorbeeld objectbeschrijvingen en contextuele informatie over erfgoedobjecten. Deze gegevens worden opgenomen in collectieinformatiesystemen om de objecten goed te kunnen beheren of toegankelijk te maken.
3. **Born digital erfgoed:** Erfgoedmateriaal dat van origine digitaal is, zoals foto's die met een digitale camera zijn gemaakt, digitale documenten, digitale kunst, websites, apps of games.

Digitaal erfgoed kenmerkt zich door unieke ‘eigenaardigheden’.¹⁶ Deze bijzondere karakteristieken maken dit erfgoed zeer kwetsbaar. Door de snelle verandering van techniek, kenmerkend voor deze tijd, dreigt op elk niveau van het digitale object het gevaar van verlies. Een digitaal object is een samenspel tussen de bitstream (de nullen en enen), de software en de hardware. Werkend zijn ze precies op elkaar afgestemd en vormen de delen een betekenisvol digitaal geheel. Doordat de techniek voortdurend verandert moet je de bitstream, de software en de hardware-omgeving 'levend' houden om het digitale object te behouden. De bitstream moet dus worden aangepast aan de nieuwe technologie (bijvoorbeeld door migratie) of de nieuwe technologie steeds van software voorzien die oude computers kan nabootsen (emulatie). Doe je dat niet, dan gaat het digitale object verloren, omdat de digitale sporen niet meer tot een betekenisvol

¹⁵ Zie 'Digitaal erfgoed' in 'ABC Digitaal Erfgoed' van DEN, het nationale kennisinstituut voor digitalisering in de cultuursector: <http://www.den.nl/abc/Digitaal-erfgoed/>

¹⁶ Om de basisbeginselen van ICT te begrijpen raden we de twee volgende inspirerende online cursussen aan. De eerste is 'Hoe werkt een computer?'. Deze mini-cursus is gemaakt in opdracht van Nemo Kennislink en bestaat uit drie delen. [Deel 1: Het hart van de computer](#). [Deel 2: de taal van de computer](#). [Deel 3: Communiceren, van muis tot internet](#). De tweede cursus is 'Computer Science', gemaakt door Crash Course. Dit is een educatief YouTube-kanaal dat is gestart door de Greenbroers, Hank Green en John Green. De presentatrice is Carrie Anne Philbin, hoofd Educatie van de Raspberry Pi Foundation. In haar introductie legt ze uit dat computers behoorlijk ingewikkeld kunnen lijken. Maar, zegt ze, eigenlijk zijn het "eenvoudige machines die complexe acties uitvoeren door vele lagen van abstractie". In deze serie neemt Philbin de kijker mee door de verschillende lagen.

object kunnen worden gereconstrueerd. Je moet digitaal materiaal dus continu in de gaten houden. Is het bestand nog leesbaar, wordt het bestandsformaat nog ondersteund, hebben we nog de juiste hardware? In Computer Science wordt ook wel gesproken over ‘abstractieniveaus’.¹⁷ De gelaagdheid van de techniek, de onderlinge afhankelijkheden en de snelle veranderingen maakt dat het digitaal erfgoed complex is. Door te denken in verschillende abstractielagen kan de complexiteit, en daarmee de kwetsbaarheid, worden begrepen. In de komende hoofdstukken gaan we hier verder op in. Het verwerven, reconstrueren, conserveren, ontsluiten en duurzaam toegankelijk houden van digitale objecten stelt dus compleet andere eisen dan fysieke objecten dat doen. Als we geen actie ondernemen, verdwijnt een deel van ons erfgoed.

Het opgraven van het digitaal erfgoed uit de vroege internet jaren voegt een extra dimensie toe aan het proces. Niet alleen is de gebruikte techniek nog meer verouderd, ook de kennis erover én de herinneringen eraan vervagen. Daarnaast speelt een heel praktisch analoog gevaar: Mensen ruimen hun zolders op en gooien harde schijven, floppies en andere mediadragers met data weg.

Naast het opgraven van digitale objecten is het ook van groot belang om de context van de objecten (digitaal én analoog) in beeld te krijgen. De contextinformatie biedt inzicht in de werking, het gebruik, de herkomst en de betekenis van het object en helpt bij het reconstrueren en duiden van het object.

¹⁷ <https://www.cise.ufl.edu/~mssz/CompOrg/CDAintro.html>

Onze aanpak: case study 'DDS Herleeft'

In 2010 maakte de afdeling E-cultuur van het Amsterdam Museum een plan om de museumwebsite 'klaar te maken voor de toekomst'. In het plan presenteerden we de nieuwe mogelijkheden van het web: Hoe bevorder je online participatie en hoe maak je je collectie online toegankelijk en deelbaar? Dit inspireerde de afdeling Collectie. Het web opende een nieuw arsenaal aan digitale objecten en nieuwe mogelijke benaderingen. Daarmee rees de vraag: Met welke nieuwe digitale objecten kunnen we de geschiedenis (en het bedreigde erfgoed) tot leven brengen en het verhaal van Amsterdam vertellen? Ons antwoord was: Graaf De Digitale Stad op en neem deze op in de collectie!

Met deze missie voorhanden formuleerden we onze eerste doelstellingen:

- Breng het internet-historisch monument DDS weer tot leven;
- Start een pilot voor webarcheologie: Hoe kun je internetgeschiedenis reconstrueren, conserveren, ontsluiten en toegankelijk maken op wetenschappelijk en maatschappelijk niveau?

Om dit te bereiken maakten we een eerste plan van aanpak. Dit is een eerste, behoorlijk 'nerdy', versie. In het volgende hoofdstuk vind je het plan van aanpak dat hieruit voortgekomen is. De eerste versie:

1. Lanceer een Open Geschiedenis Lab en Virtueel Museum
2. Bring Out Your Hardware & Finding Lost Data ('The Grave Diggers Party')
3. The Rise of the Zombies: Analyseer en maak een reconstructie
4. The Flight of the Zombies: Presenteer de Lost & Found data en de reconstructie (DDS in 'multiple-points-in-time')
5. Enlightenment: Let the Bytes Free! Conclusies en evaluatie

In de volgende hoofdstukken laten we je zien hoe je zelf een webarcheologisch project kan starten. We bieden praktische handvaten en vervolgens illustreren wij deze, waar mogelijk, aan de hand van onze case study van DDS.

05

Plannen

Hoe start je een webarcheologisch project?

In dit hoofdstuk helpen we je bij het schrijven van een plan en tref je de voorbereidende werkzaamheden voor jouw digitaal-erfgoed project. Je schrijft een plan om je doel te bereiken. Met een goed plan is het goed overtuigen. Vaak is het plan de eerste gelegenheid voor directie, medewerkers en partners om kennis te nemen van jouw project. Houd het plan kort en bondig, maar wees niet bang om anekdotes en citaten op te nemen. Hierdoor gaat het project leven. Je gaat aan de slag met een nieuw vakgebied, waarbij je een of meerdere relatief jonge (born digital) objecten, die nog niet zo lang geleden verloren zijn gegaan, gaat opgraven en reconstrueren met relatief nieuwe (digitale) gereedschappen. Je plan valt hoogstwaarschijnlijk nog buiten de scope of buiten het collectiebeleid van je organisatie. Wees je hiervan bewust en zorg dat je goed beslagen ten ijs komt. Overtuig je collega's met je plan van de waarde en de betekenis van deze nieuwe vormen van erfgoed voor je organisatie.

Stappenplan:

- 1 Beschrijf de doelstellingen
- 2 Doe bronnenonderzoek
- 3 Beschrijf het digitale object en het belang ervan
- 4 Breng in kaart wat je nodig hebt
- 5 Maak een plan van aanpak
- 6 Stel een begroting op
- 7 Zet een communicatieplatform op

5.1 Beschrijf de doelstellingen

Het lastige van een webarcheologisch project (net als een archeologische opgraving) is dat je de uitkomst niet van tevoren weet. Je weet bijvoorbeeld niet of je het object daadwerkelijk kan 'opgraven', wat je precies zult vinden en in welke staat het born digital materiaal verkeert. Op vele vlakken in het proces (graven, reconstrueren, conserveren, archiveren, ontsluiten en toegankelijk maken) kan er een kink in de kabel komen. De haalbaarheid kun je dus pas achteraf bepalen. Maar wees niet te voorzichtig in het beschrijven van je doelstellingen. Wees ambitieus. En geef waar mogelijk een reële inschatting van de haalbaarheid per doel. Denk bijvoorbeeld aan de volgende vragen: Wat wil je opgraven en behouden? Waarom wil je het opgraven? Voor wie graaf je het op? Wie gaat wat doen? Waar gebeurt wat? Wanneer gebeurt wat?

Onze aanpak: case study 'DDS Herleeft'

Zoals beschreven in hoofdstuk vier 'Getting started' hadden wij al een missie en een eerste plan van aanpak geformuleerd. Op basis hiervan werkten wij onze doelstellingen uit. Door voortschrijdend inzicht hebben we deze inmiddels aangescherpt.

- Behoud het internet-historisch monument DDS;
- Geef DDS de plek die het verdient, binnen de geschiedenis van Amsterdam en in de collecties van onze erfgoedinstellingen;
- Vergroot de bewustwording voor het gevaar van 'digitaal geheugenverlies';
- Bied kleinere musea en organisaties inzicht in de specialistische kennis over het reconstrueren en duurzaam behoud van born digital erfgoed en verlaag de drempel voor toekomstige webarcheologische projecten;
- Bied inzicht in de (bestaande en nieuwe) processen, technieken en methoden om born digital materiaal en de context waarin zij worden aangetroffen, op te graven en te reconstrueren;
- Maak DDS data waar mogelijk 'future-proof', zodat deze voor zoveel mogelijk geïnteresseerden toegankelijk zijn en het materiaal in de toekomst (her)gebruikt kan worden:¹⁸
 - Zichtbaar (content): om (her)gebruik van DDS collectie te bevorderen.
 - Bruikbaar (verbinding): om gebruik van DDS collectie te verbeteren door het beschikbaar te stellen, door data te verbinden en te verrijken.
 - Houdbaar (diensten): om DDS duurzaam te behouden en toegankelijk te maken.

¹⁸ Zie: <https://www.rijksoverheid.nl/documenten/publicaties/2015/03/09/nationale-strategie-digitaal-erfgoed> en <https://www.dtls.nl/fair-data/fair-data/>.

TIP: Zet een strategie op voor een duurzame digitale collectie

Toen het idee in 2010 ontstond om De Digitale Stad op te graven realiseerden we ons wat we ons op de hals haalden. Het museum had al een online collectie, een actieve rol op social media en een aantal actieve online communities. Maar een digitaal duurzaamheidsbeleid hadden we nog niet. Eerst doen, dan beleid maken was ons uitgangspunt.

Wil jij echter eerst aan de slag met het schrijven van een beleidsplan over het vormen van een duurzame digitale collectie met born digital materiaal? Laat je dan inspireren door “Het scoremodel”¹⁹ en de “Catalogus Duurzaamheidsbeleid”.²⁰ Het scoremodel is een online vragenlijst die aan het einde een rapport oplevert waarin de grootste risico's voor jouw organisatie op een rij zijn gezet. Het is een tool voor kleine en middelgrote organisaties die een 'digitale collectie' beheren en deze voor lange tijd willen bewaren. Maar ook particulieren of grote organisaties kunnen er mee aan de slag. De 'Catalogus Duurzaamheidsbeleid' levert bouwstenen voor het opstellen van een duurzaamheidsbeleid en is een hulpmiddel voor organisaties bij het ontwikkelen van duurzaamheidsstrategieën en de implementatie ervan. De catalogus richt zich op (medewerkers van) collectiebeherende instellingen.

In een duurzaamheidsbeleid beschrijf je je doelstellingen met betrekking tot het bewaren van duurzame toegankelijkheid van de materialen die in je digitale archief zijn opgeslagen.²¹ Het gaat hier nadrukkelijk niet om het hosten van je digitale collectie, maar juist om de duurzame digitale toegankelijkheid ervan. Je kunt bijvoorbeeld een infrastructuur voor langetermijnopslag ('digitaal depot' of e-depot) inrichten of inhuren, of de data veilig stellen, bijvoorbeeld door te kiezen voor een van de duurzaamheidsstrategieën: migratie, conversie, emulatie of herinterpretatie van de data. Of je kiest voor meerdere opties, zoals in onze case study (zie verderop in het DIY Handboek). Migratie is het verhuizen van de data naar een ander opslagmedium, conversie is het overzetten naar een ander bestandsformaat, emulatie betekent het nabootsen van een verouderde soft- en hardware-omgeving zodat oude bestanden toch gebruikt kunnen blijven worden en herinterpretatie behelst het opnieuw installeren, uitvoeren of reconstrueren van een digitaal object met nieuwe technieken.²²

¹⁹ <https://www.scoremodel.org/>

²⁰ Bekijk de 'Catalogus Duurzaamheidsbeleid' op de wiki Duurzame toegang tot digitale informatie, opgezet en ingericht door de Nationale Coalitie Digitale Duurzaamheid (NCDD) in samenwerking met het Netwerk Digitaal Erfgoed, <http://wiki.ncdd.nl/>

²¹ "Duurzaamheidsbeleid" door Barbara Sierman en Marcel Ras, OD (mei-juni 2017), <http://www.ncdd.nl/news/artikel-duurzaamheidsbeleid-od-magazine/>

²² "Generieke workflows born digital erfgoed. Behoud van born digital erfgoed in Nederland: film, fotografie, architectuur, kunst". Onderzoek van de Culturele Coalitie Digitale Duurzaamheid naar born digital erfgoed, door Gaby Wijers en Hannah Bosma. CCDD (2015). http://ncdd.nl/site/wp-content/uploads/2015/10/CCDD-BornDigitalOnderzoek-def_2015.pdf

5.2 Doe bronnenonderzoek

De volgende stap is het doen van bronnenonderzoek, waardoor je zoveel mogelijk te weten komt over je gewenste digitale object en de context. Begin met de bekende W-vragen: wie, wat, waar, wanneer en waarom.

Online bronnen

Steeds meer bronnen verschijnen online. Het web is een goudmijn voor iedere onderzoeker. De Wayback Machine van het Internet Archive noemden we al. Het Internet Archive is een digitale bibliotheek van internetsites en andere culturele artefacten in digitale vorm. In Delpher²³ vind je miljoenen gedigitaliseerde teksten uit Nederlandse kranten, boeken en tijdschriften die je makkelijk kunt doorzoeken. De teksten komen uit de collecties van diverse wetenschappelijke instellingen, bibliotheken en erfgoedinstellingen. De teksten zijn vrij te gebruiken voor privédoeleinden en onderzoek. Zoek en bekijk ook andere (internationale) krantenbanken.²⁴ Een andere interessante bron is Memento.²⁵ Dit is een zoekmachine voor webarchieven. Je doorzoekt alle aangesloten webarchieven in één klik.

Onze aanpak: case study 'DDS Herleeft'

Voor het schrijven van ons plan hebben we gebruik gemaakt van traditioneel historisch bronnenonderzoek. Er is in de loop van de jaren veel (academisch) onderzoek gedaan naar de DDS. Dit Handboek put uit deze rijke bron van publicaties. Een aantal publicaties staan als referentie in het hoofdstuk 'Bronnen'.

Door het nieuwe object van onze studie (de virtuele stad als digitaal erfgoed object) en onze nieuwe onderzoeksvragen zijn we vooral op zoek gegaan naar nieuwe (vooral digitale) bronnen. De Wayback Machine van het Internet Archive bood ons een unieke rijke bron aan digitale sporen. Via de Wayback Machine zijn verschillende tijdsopnames ('captures'), historische varianten, van de DDS-site en van de DDS-huizen, te bekijken.²⁶ De eerste digitale sporen dateren van 2 november 1996. DDS was toen al ruim tweeëneenhalf jaar online. Tussen 2 november 1996 en nu (22 augustus 2017) is <http://www.dds.nl> 9.449 keer (statisch) gearcheveerd, en <http://huizen.dds.nl/> 6.833 keer. Op 15 mei 2001 zijn de huizen voor de laatste keer gearcheveerd. Hierna verwees huizen.dds.nl door naar www.dds.nl, de website van DDS als provider. Ons onderzoek richtte zich alleen op de periode dat de virtuele stad publiek domein was. Naast de Wayback Machine waren voor ons ook de Usenet nieuwsgroepen een rijke bron van informatie. Usenet bestond al voor [www](http://www.usenet.nl) en is dankzij Dejanews compleet gearcheveerd. Daarna heeft Google de nieuwsgroepen overgenomen en deze zijn nu live te benaderen.²⁷ Ook het archief van Nettime-nl was een goede bron. Nettime-nl is een mailinglist over netkritiek, media-activisme en netart. De lijst is in 1996 opgezet en vormt een rijke bron over DDS.²⁸

5.3 Beschrijf het digitale object en het belang ervan

Na het bronnenonderzoek ga je het digitale object beschrijven. Als je in een collectiehoudende instelling werkt, kun je je bij de beschrijving laten inspireren door acquisitie formulieren voor de collectie. Hierin staan de belangrijke argumenten op basis waarvan beslissingen genomen worden om een object in de collectie op te nemen.²⁹ Denk hierbij aan de waarden die het object vertegenwoordigt, zoals bijvoorbeeld de artistieke of technische waarde, uniciteit en belevingswaarde, de herkomstgeschiedenis en de eigenaar en/of contactpersoon voor meer informatie. Beschrijf je motivatie: Waarom is het belangrijk om dit object op te graven, te reconstrueren en op te nemen in de collectie? Beschrijf ook hoe dit object zich verhoudt tot de missie en visie van jouw organisatie. Lever bij je beschrijving alle relevante documentatie aan.

TIP: Van wie is het digitale object?

Wij kenden de initiatiefnemers en betrokkenen door de jaren heen van DDS, maar als je dat niet weet, gebruik dan de tool 'Whois'. Met deze tool kun je meer informatie krijgen over een domeinnaam, zoals de eigenaar van een website of de provider die de naam geregistreerd heeft. Een erg handige tool voor webarcheologen!³⁰ Whois geeft echter alleen de huidige informatie. De oude informatie kan ook gevonden worden via de 'Whois History' op websites zoals <http://whois.domaintools.com>.

TIP: Bewaar je bronnen

Bewaar vanaf het begin van je project alle gevonden online artikelen (webpagina's). Er bestaat een grote kans dat artikelen verdwijnen (websites verhuizen, worden opgeschoond of opgeheven, nieuwe domeinnamen worden gebruikt, de inhoud van een site wordt achter een betaalmuur geplaatst etc.). Let wel op als je gebruik maakt van social-bookmarking websites. Ook deze kunnen verdwijnen. Zoals bijvoorbeeld del.icio.us.³¹ Deze site is inmiddels zes keer van eigenaar veranderd en wordt nu vervangen door een betaalde dienst. Maak altijd een back-up! Maak screenshots van informatie. Scan al je gevonden papieren stukken vanaf het begin met OCR (Optical Character Recognition), zodat de stukken doorzoekbaar (en vindbaar) zijn. Doe dit meteen, want doe je dit later, dan stapelt het werk zich snel op.

²³ <http://www.delpher.nl/>

²⁴ https://en.wikipedia.org/wiki/Wikipedia:List_of_online_newspaper_archives

²⁵ <http://timetravel.mementoweb.org/>

²⁶ Bekijk verschillende versies van de DDS site: http://web.archive.org/web/*/http://www.dds.nl en de historische DDS huizen: http://web.archive.org/web/*/http://huizen.dds.nl/

²⁷ <https://groups.google.com/>

²⁸ <https://nettime.org/archives.php> (scroll door naar nettime-nl: Nettime Nederlands).

²⁹ Lees ook: Richtlijnen voor de selectie van digitaal erfgoed voor langetermijnconservering, door: UNESCO/PERSIST Content Task Force (2016). https://www.unesco.nl/sites/default/files/uploads/Comm_Info/finalpersistcontentguidelinesfinal_dutch.pdf

³⁰ Kijk bijvoorbeeld bij <https://www.sidn.nl/whois>

³¹ Deze gebruikten wij ook voor ons project: <https://del.icio.us/re DDS>.

Onze aanpak: case study 'DDS Herleeft'

Wij omschreven het belang van ons digitale object, De Digitale Stad, als volgt in ons plan.

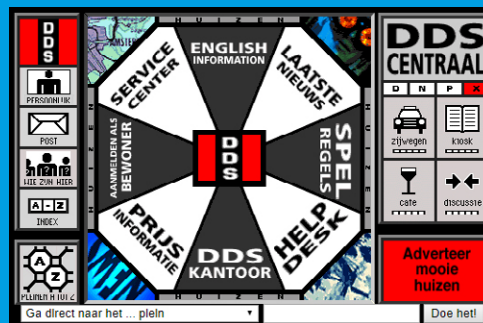
DDS is de oudste Nederlandse online gemeenschap en speelde een belangrijke rol in de internetgeschiedenis van Amsterdam en Nederland. Voor het eerst was internet (gratis) toegankelijk voor het grote publiek in Nederland. Voor veel mensen betekende de virtuele stad in 1994 de eerste kennismaking met de mogelijkheden van het internet. De virtuele stad en haar inwoners brachten objecten, ideeën en tradities voort in nieuwe digitale (hybride) vormen, zoals webpagina's, nieuwsgroepen, chats, audio en video. DDS laat zien hoe bewoners zich presenteerden op het vroege web en hoe ze met elkaar omgingen. DDS is een belangrijke en unieke historische bron over de beginjaren van het internet in Nederland en een schatkamer van digitaal erfgoed. Het fungeerde ook als een proeftuin en opereerde op het snijvlak van kunst, cultuur, technologie en wetenschap. Naast een experiment met computers, was het een experiment met vraagstukken en uitdagingen uit het nieuwe opkomende informatietijdperk.

DDS is een digitale omgeving met cultuurhistorische en technologische lagen. De virtuele stad was niet alleen een website, maar een complex, dynamisch informatiesysteem met verschillende applicaties door de tijd heen. DDS maakte veelal gebruik van experimentele, hybride, niet-standaard technologie. DDS volgde de nieuwe ontwikkelingen altijd op de voet, en dat resulteerde in een aantal unieke 'stadsgezichten' (interfaces, verschijningsvormen):

- 1994-1995: DDS 1.0 - 15 januari 1994; telnet-interface. Alle informatie en communicatie werd aangeboden in de vorm van een op tekst gebaseerde omgeving met tekst-menu's (command-line interface, MS-DOS, UNIX) in Bulletin Board System-technologie. De zogenaamde 'Free-Nets' in de Verenigde Staten en Canada waren een belangrijke bron van inspiratie voor de oprichters. Free-Nets waren 'community netwerken', of 'virtuele gemeenschappen', ontwikkeld en uitgerold door burgers zelf: burgernetwerken. De metafoor van de stad kwam tot uiting in de organisatie van de interface. Er was een postkantoor (voor e-mail), openbare forums om andere bezoekers te ontmoeten, een stadhuis en een centraal station (de toegangspoort tot het internet). Door de metafoor van de stad werd het web een plaats waar je kon intrekken en leven.



DDS 1.0



DDS 3.0

- 1994-1995: DDS 2.0 - 15 oktober 1994; website: statische HTML. Overstap naar het World Wide Web (www). Eerste DDS website met een grafische interface, hyperlinks, tekst en plaatjes.
- 1995-2001: DDS 3.0 - 10 juni 1995; websysteem: statische HTML en interactieve webpagina's. Alle informatie en communicatie werd geïntegreerd aangeboden vanuit één nieuwe interface: de meest bekende, pleinen-interface. Elk plein heeft een eigen thema en karakter en diende als ontmoetingsplaats voor mensen die geïnteresseerd waren in dat specifieke thema. Bezoekers konden informatie en ideeën uitwisselen met elkaar. 'Bewoners' konden een eigen 'huis' bouwen (een webpagina), e-mails verzenden en ontvangen (wereldwijd!), deelnemen aan discussiegroepen, chatten in cafés, deelnemen aan de 'Metro', stemmen etc.

De Digitale Stad is, in al haar verschijningsvormen, belangrijk erfgoed voor Amsterdam. DDS markeerde een punt in de geschiedenis, was iets typisch Amsterdams, maakte de ontwikkeling van internet voelbaar en illustreerde de rol van Amsterdam in de internet revolutie. Hieronder hebben wij citaten opgenomen uit onze beschrijving over ons (digitale) object DDS om de geschiedenis invoelbaar te maken.

Voor het eerst werd het internet in Nederland toegankelijk voor het grote publiek. Voor veel mensen betekende DDS in 1994 de eerste kennismaking met de mogelijkheden van internet. DDS was ook een van de weinige websites in het Nederlands wat samen met de stadsmetafoor zorgde voor laagdrempelige toegang. Inbellen, e-mail en een eigen webpagina was dus gratis, een unicum in die tijd.

"De initiatiefnemers, een club jonge computerfreaks, dromen al tijden van een groot – publiek toegankelijk – netwerk. Zij zien het als de enige manier om de maatschappelijke consequenties van moderne technologie voor een breed publiek duidelijk te maken. 'Wij willen mensen op het Internet brengen die er anders nooit zouden komen,' stelt Marleen Stikker, medewerkster van De Balie. 'En op die manier een debat op gang brengen over de mogelijkheden van computernetwerken!'"³²

DDS was onmiddellijk een groot succes. Al snel waren alle modems in Amsterdam en omstreken uitverkocht.³³ Binnen tien weken schreven zich 12.000 bewoners in. DDS was zelfs zo succesvol dat er files bij de digitale poorten ontstonden. In de loop van de jaren groeide het aantal bewoners flink: in 1994 waren er 12.000 gebruikers, in 1995: 33.000, 1997: 60.000, 1998: 80.000 en in 2000: 140.000.

"Opeens konden duizenden mensen op een laagdrempelige manier kennismaken met de mogelijkheden van internet. In Amsterdam was nauwelijks nog een modem te krijgen. De inbellijnen van DDS waren permanent bezet. Wie contact wilde maken met de stad, moest vaak tientallen keren opnieuw inbellen. 'We hadden één modembankje met dertig modems', herinnert Joost Flint zich. 'Vaak stonden er honderd mensen in de rij om binnen te komen.' Wie zelf geen computer met modem had, kon bij vijf openbare gelegenheden in Amsterdam inloggen. In het politiek-cultureel centrum De Balie (één van de initiatiefnemers van DDS), de openbare bibliotheek, het stadhuis en het Stedelijk Museum stonden computers waarmee mensen dertig minuten gebruik konden maken van DDS. Na dertig minuten werd de verbinding verbroken en was iemand anders aan de beurt."³⁴

Voor het eerst zetten organisaties voet op het internet. Onder andere de Openbare Bibliotheek, de gemeente Amsterdam, het Anne Frank Huis, kranten als NRC Handelsblad en Het Financiële Dagblad begonnen hun digitale bestaan bij DDS.³⁵

³² "De Digitale Stad van start" door Francisco van Jole, de Volkskrant (15 januari 1994), <https://fjole.home.xs4all.nl/archief/artikelen/Volkskrant/1994/dds.html>

³³ "10 jaar De Digitale Stad: Debat Publiek Domein", door Waag Society, Indymedia.nl (14 januari 2004), <https://www.indymedia.nl/nl/2004/01/16343.shtml>

³⁴ "Nog altijd dorpsruzie bij jubilerend DDS", door Maarten Reijnders, Webwereld.nl (13 januari 2004), <http://webwereld.nl/e-commerce/2565-nog-altijd-dorpsruzie-bij-jubilerend-dds>

³⁵ "Lot in eigen hand", door Marie-José Klaver, NRC Webpagina's (26 januari 2001), <http://retro.nrc.nl/W2/Lab/At/20010126.html>

De opening van DDS was een typisch Amsterdams fenomeen:

"Vanavond opent de Digitale Stad haar elektronische poorten. Dat lijkt wel de eerste zin van een sprookje en eigenlijk is het dat ook een beetje. Door een monsterverbond van anarchistische hackers en autoriteiten is in Amsterdam binnen drie maanden een openbaar elektronisch communicatiesysteem uit de grond gestampt dat zijn weerga op het Europese vasteland niet kent." (...) "Cultureel centrum De Balie en de Stichting Hacktic Netwerk, de bouwers van de stad, hebben de gemeente warm weten te krijgen voor hun ideaal: een open communicatiesysteem voor burgers. In het zicht van de gemeenteraadsverkiezingen voelden de politici daar wel wat voor. Als die achter de rug zijn, moet de Digitale Stad op 1 april aanstaande haar poorten weer sluiten. Of er moet in de tussentijd een andere geldschieter of financieringsmogelijkheid gevonden worden." (...) "De Digitale Stad is niet zomaar een project, het is een politieke daad."³⁶

DDS maakt de ontwikkeling van internet voelbaar:

"Kennismaken met Internet veroorzaakt een merkwaardige sensatie. Zo'n gevoel dat zich meester van je maakt als je voor het eerst de St Pieter in Rome betreedt of een bezoek brengt aan de Niagara-watervallen. Alleen is het bij Internet niet de pracht en praal of het natuurschoon dat imponeert maar de overdonderende hoeveelheid informatie en mogelijkheden."³⁷

Twee herinneringen van bewoners van het eerste uur:

Michaël van Eeden, programmeur, vertelt over De Digitale Stad:

"Het zou een plek op internet zijn die eruitzag als een stad, waarin bewoners met elkaar konden praten over gezamenlijke interesses en konden e-mailen met de rest van de wereld. Er was een museum, een bibliotheek, een stadhuis en een postkantoor. Dat is het, dacht ik, en ik besloot dat ik daar bij wilde zijn. Het was ongelooflijk spannend. Ik ging erheen, het was een bitter koude winter. DDS zat in een leeg pand van de scheepvaartschool. Via de wc van een naastgelegen café was het te betreden. Daar, in een enkele kamer, zaten dag en nacht mensen bij elkaar om een paar computers heen. Op de grond lagen kampeermatrasjes. Enkele tientallen modems flikkerden. Die lampjes betekenden dat er mensen in de digitale stad waren. Ik vond dat heel wat. Er gebeurde iets belangrijks, dat voelde je."³⁸

³⁶ "De Digitale Stad van start", door Francisco van Jole, de Volkskrant (15 januari 1994), <https://fvjole.home.xs4all.nl/archief/artikelen/Volkskrant/1994/dds.html>

³⁷ "Internet is een nieuw werelddeel", door Francisco van Jole, de Volkskrant (1 mei 1993), <https://fvjole.home.xs4all.nl/archief/artikelen/Volkskrant/1993/internet.html>

³⁸ "Red de stad!" door Romana Abels, Trouw (6 april 2001), <https://www.trouw.nl/home/red-de-stad~a1909c66/>

Marianne van den Boomen, auteur en docent onderzoeker:

*"Ik had op dat hackerskamp in de zomer van 1993 (HEU, red) internet voor het eerst gezien en had echt het idee: dit gaat de wereld veranderen, en ik wil erbij zijn. Maar makkelijk was het niet. Op het kamp, te midden van jongetjes die me met alles hielpen kon ik wel uit de voeten, maar later zat ik thuis tegen een groen dollarteken op een verder leeg scherm aan te kijken. Zo zag het internet er toen uit: geen web, geen plaatjes, alles moest met ondoorzichtige commando's."*³⁹

DDS illustreert de rol van Amsterdam in de internetrevolutie. Door het gebruik van de stadsmetafoor, de aanpak en het succes trok DDS veel aandacht in het buitenland. DDS speelde een belangrijke rol in de cyberreputatie van de stad Amsterdam.

*"Met stijgende opwindning had Reineke van Meerten (destijds projectmanager Telematica Stimuleringsbeleid van de gemeente Amsterdam, red.) geluisterd naar de jongelui die aan haar bureau waren verschenen. Zij schetsten voor haar De Digitale Stad, een computernetwerk, aangesloten op een wereldwijd 'iets' dat 'internet' heette en waarin Amsterdam zou worden omgetoverd tot een virtuele stad. Op elektronische 'pleinen' zouden Amsterdammers elkaar en hun bestuurders opzoeken om te discussiëren over actuele vraagstukken en via 'postkantoren' zou in 'e-mails' vrijelijk en razendsnel informatie verspreid kunnen worden. Onafhankelijk van plaats, tijd en sociale status. In die 'cyberspace' zou de vermaledijde kloof tussen bestuur en burgers kunnen worden geslecht. Het project De Digitale Stad beantwoordde aan het gemeentelijke beleidsdoel om bestuurlijke vernieuwing te relateren aan moderne media. De subsidieaanvraag voor het inrichten van De Digitale Stad (DDS) werd tot tweemaal toe gehonoreerd met 100 duizend gulden." (...) "DDS werd zelfs wereldnieuws. Van Meerten: 'We gingen op presentatietoernee. Ik herinner me er een bij de Raad voor Europa en een bij een ander Europees overleg in Parijs. Zelfs uit Washington kwam een uitnodiging. Iedereen dacht dat de 'Digital City' van de gemeente was. We hebben er internationaal zeer goede sier mee gemaakt."*⁴⁰

³⁹ "Red de stad!" door Romana Abels, Trouw (6 april 2001), <https://www.trouw.nl/home/red-de-stad-~a1909c66/>

⁴⁰ "Elektronische overbrugging van 'de kloof'", BMC, Jaargang 6, nummer 2 (2006), <https://hart.amsterdam/nl/page/33440/project-re-dds-ii-waarom-dds>

De cyberreputatie van Amsterdam werd ook door CNN opgemerkt:

*"Amsterdam is a city building on it's rich past to create a high tech future. For hundreds of years the city of Amsterdam has been a center of commercial trade, art and education. Now it's helping point the way in the information revolution too. For 2 reasons. Holland has the highest rate of new internet connections in Europe, and free or low cost internet access and e-mail service for everyone who wants it."*⁴¹

5.4 Breng in kaart wat je nodig hebt

Door de kennisgebieden en middelen (methoden, mensen en geld etc.) in kaart te brengen leer je wat jouw organisatie in huis heeft en wat je moet (kunt) uitbesteden.

Stappenplan:

- Breng de kennisgebieden in kaart
- Doe vooronderzoek naar bestaande methoden
- Maak een overzicht van de juridische en ethische aspecten
- Vind de juiste mensen

Breng de kennisgebieden in kaart

Als webarcheoloog krijg je te maken met bekende en onbekende vakgebieden, waaronder collectiebeheer en -ontsluiting specifiek voor born digital materiaal, digitaal archiveren, wetgeving en ethiek, communitymanagement, projectmanagement, fondsenwerving, communicatie en techniek. Jij moet uitzoeken wat je zelf kunt en wat je moet uitbesteden. We lichten een aantal velden toe in deze paragraaf en in de komende hoofdstukken.

⁴¹ "Digital Expansion in The Netherlands". Verslag uit Amsterdam in het programma "The CNN computer connection" door CNN reporter Dick Wilson, CNN (16 juni 1997), <https://hart.amsterdam/nl/page/33440/project-re-dds-ii-waarom-dds>

Onze aanpak: case study 'DDS Herleeft'

Wij zijn begonnen met het in kaart brengen van de benodigde kennisgebieden, die we kort omschreven in een overzicht met de naam '23 Dingen Webarcheologie':⁴²

1. Digitale archeologie
2. Born digital materiaal
3. Webarchivering (webscraping/ webharvesting)
4. Hardware archiving (verouderde hardware, firmware/BIOS)
5. Crowdsourcing
6. Crowdfunding
7. Hacking (hackers cultuur, reverse engineering, beveiliging)
8. Lock-in (closed formats, closed source, DRM, encrypted back-ups, key dongels, license servers, propriety software, NIST Software Archive, authoritative information, externe dependencies)
9. Juridische en ethische kwesties
10. Emulatie
11. Formaten (bestandsformaten, archiefformaten)
12. Metadata
13. Software
14. Beleid voor musea (beheer, acquisitie, etc.)
15. AV-archivering
16. Digitale tijd-capsules
17. Semantisch web
18. Aanbevelingen voor programmeurs en makers (open data, Creative Commons, open source, source code, full disk back-ups, versioning control)
19. Werkomgevingen (locaties)
20. Presentatie
21. Storytelling
22. Internationalisering
23. Party: vieren van successen!

Deze lijst heeft ons geholpen de juiste experts te vinden. Onder de vlag van '23 Dingen Webarcheologie' hebben we ons eerste netwerk van experts opgezet en verschillende bijeenkomsten georganiseerd. Door dit netwerk hebben wij onze huidige projectpartners gevonden.

⁴² We lieten ons inspireren door het concept van '23 dingen'. Deze vorm hielp ons om de mogelijke aspecten van webarcheologie te belichten. '23 dingen' was gebaseerd op het voorbeeld van het Amerikaanse 23 Things van Helene Blowers (Public Library of Charlotte & Mecklenburg County). Zie: <http://plcmclearning.blogspot.nl/>

TIP: Breng locaties in kaart

Breng alvast locaties in kaart waar je samen kunt komen met deelnemers om te 'graven' en reconstrueren. Beter op tijd bedacht dan te laat geboekt en misgeschoten. Denk aan locaties met de benodigde apparatuur waar je mogelijk gebruik van kunt maken, internetverbinding en een goed koffiezetapparaat. Voorbeelden zijn universiteiten, bibliotheken, musea, medialabs, fablabs en hackerspaces.

Doe vooronderzoek naar bestaande methoden

Nu je meer weet over je kennisgebieden kun je beginnen met een vooronderzoek. Kijk om je heen naar bestaande methoden en wat andere webarcheologen doen. Begin met een desktop onderzoek. Wat gebeurt er om je heen? Zijn er voorbeelden die jou kunnen helpen? Zijn er mensen met wie je contact kunt opnemen?

Onze aanpak: case study 'DDS Herleeft'

We vonden de volgende internationale projecten die ons inspireerden. We maakten een onderscheid tussen webharvestingprojecten en webarcheologische projecten. Bij webharvesting projecten worden actuele (live) data geoogst (ook wel webharvesten, scrapen of mirroren genoemd) en weergegeven of gevisualiseerd. Bij webarcheologische projecten worden de digitale objecten, de originele data en software, die niet meer online zijn (het 'dode' web) weer 'opgegraven', gereconstrueerd en op de oorspronkelijke hardware of via emulatie technieken getoond. Hierdoor kan het publiek de oorspronkelijke werkende website bekijken en weer ervaren. Bekijk de voorbeelden.

Webharvesting-projecten (statisch webarchiveren):

- De Wayback Machine van het Internet Archive.⁴⁴ Dit is het eerste digitale archief dat sinds 1 januari 1996 webpagina's archiveert en ontsluit via een webinterface.⁴⁵
- GeoCities: GeoCities ontstond eind 1994 in Californië en bood gebruikers, net als DDS, de mogelijkheid om gratis een homepage te maken. In oktober 2009 trok Yahoo, eigenaar van GeoCities, de stekker uit deze site. Verschillende organisaties hebben zich ingezet om de data te redden.⁴⁶

⁴⁴ <http://archive.org/web/>

⁴⁵ The Wayback Machine: Preserving the History of Web Pages door ForaTv: <https://www.youtube.com/watch?v=Jsl1TADosN0>. The Internet Archive Documentary, door Doc Streams in 2014: <https://www.youtube.com/watch?v=1WxGDOeS-s4>. Bekijk oude versies van de DDS site in de Wayback Machine: http://web.archive.org/web/*/http://www.dds.nl. Bekijk oude DDS huizen: http://web.archive.org/web/*/http://huizen.dds.nl/

- Hyves: Nadat in november 2013 bekend werd dat de sociaalnetwerksite Hyves zou sluiten begon het Archive Team met het binnenhalen ('scrapen') van de Hyves-profielen. Het team had tot december 2013 de tijd, daarna ging Hyves offline en kreeg de site een doorstart als een spelletjesportal. 9.000.000 profielen zijn gered. Het Hyves archief is opgenomen in de collectie van het Internet Archive.⁴⁷

Webarcheologische projecten (statisch en dynamisch webarchiveren):

- Info.cern.ch: Het kernfysisch instituut CERN in Genève heeft in 2013 de eerste website ter wereld gerestaureerd. Het onderzoeksinstituut hoopt met "Restoring the first website. A project to restore info.cern.ch – the world's first website" toekomstige generaties meer te laten zien (en beleven) van het ontstaan van het internet.⁴⁸
- De 'Digital Archaeology' tentoonstellingen: Curator Jim Boulton is de trekker van verschillende tentoonstellingen onder de noemer 'Digital Archaeology'. 'Digital Archaeology' debuteerde op de Internet Week Europe 2010. In deze (en volgende) tentoonstellingen, 'Error 404' en 'Digital Revolution', wordt de glorie tijd van de website gevierd door baanbrekende, cultureel belangrijke websites in hun oude glorie terug te brengen en tentoon te stellen met de software en op de hardware van toen. Voor de tentoonstelling 'Digital Revolution' herenigde Jim Boulton de bovengenoemde eerste webpagina ter wereld info.cern.ch met de eerste browser, Nexus webbrowser, en toonde deze op een NeXTCube server.⁴⁹

Maak een overzicht van de juridische en ethische aspecten

Niet alleen de techniek van het digitale object kan als overdonderend worden ervaren, de juridische en ethische kant kan dat ook zijn.⁵⁰ De technische veelzijdigheid van digitale objecten (code, software, hardware, data) kent zijn parallel in de wetgeving. In het geval van een community website heb je bijvoorbeeld niet alleen te maken met copyright van de software, maar ook met gebruikersdata (persoonsgegevens). Dit is des te belangrijker omdat de meeste gebruikers nog in leven zijn, in tegenstelling tot de

⁴⁷ Archive Team: <http://archiveteam.org/index.php?title=Hyves>. Angering the Hyves collection van het Internet Archive: <https://archive.org/details/hyves>. Lees Webarchivering in de praktijk: Hyves: <http://hart.amsterdammuseum.nl/nl/id/37144>

⁴⁸ <http://info.cern.ch/>. <http://first-website.web.cern.ch/>. De eerste webpagina ter wereld info.cern.ch <http://www.w3.org/History/19921103-hypertext/hypertext/WWW/TheProject.html>. Lees ook over de restauratie van de eerste website uit de Verenigde Staten: "Stanford Libraries unearths the earliest U.S. website" door Gabrielle Karampelas (Stanford Report, 29 oktober 2014) <http://news.stanford.edu/news/2014/october/slac-libraries-wayback-102914.html>

⁴⁹ <http://digital-archaeology.org/category/exhibition/>. <https://www.barbican.org.uk/hire/exhibition-hire-bie/digital-revolution>. <https://www.youtube.com/user/DigitalArchaeology>. Bekijk een demonstratie van de eerste webbrowser, de Nexus browser: <https://youtu.be/3c3Rt6QbHDw>

⁵⁰ Een bron van inspiratie was het rapport 'Digital Forensics and Born-Digital Content in Cultural Heritage Collections' door Matthew G. Kirschenbaum, Richard Ovenden, Gabriela Redwine (December 2010), <https://www.clir.org/pubs/reports/pub149>

bewoners van Pompeï... Wees je bewust van het feit dat het web nog geen 30 jaar bestaat. Als webarcheoloog kun je te maken krijgen met onder andere de volgende wetgeving: eigendomsrechten, auteursrechten, de Archiefwet, Wet Bescherming Persoonsgegevens (privacy, geheimhoudingsverklaring), Merkenrecht, Patentrecht en WOB overwegingen. Het is dus verstandig om een jurist advies te vragen voordat resultaten openbaar gemaakt worden. Bespreek hierbij het volgende: Van wie is het 'opgegraven' digitale object? Wat mogen we? Wat moeten we? Welke consequenties kan dit hebben voor het verwerven, reconstrueren, conserveren, ontsluiten en duurzaam toegankelijk houden van de digitale objecten?

TIP: Handel met zorgvuldigheid

Maak een juridisch document waarin je de juridische en ethische zaken adresseert om aan te tonen dat je met gepaste zorgvuldigheid hebt gehandeld. Breng in kaart wie welk onderdeel uitvoert en waar deze onderdelen aan moeten voldoen. Documenteer het goed zodat je bewijs hebt van wat je wel en niet hebt gedaan. Hierdoor voorkom je eventuele misverstanden en mogelijke conflicten. Maak ook duidelijke afspraken over hoe materiaal wordt aangeleverd, wanneer, waarom, onder welke voorwaarden en beschrijf wat de rechten en plichten zijn voor en na de overdracht van de objecten.

Vind de juiste mensen

Gezien de urgentie, de complexiteit, aard en omvang van het onderwerp (en zoals bij ons, onbekendheid met de materie), is het zeer aannemelijk dat je de kennis van en de samenwerking met verschillende instellingen nodig hebt. Nadat je de kennisgebieden, waaronder een overzicht van de juridische en ethische aspecten, in kaart hebt gebracht kun je op zoek gaan naar de experts. En, minstens zo belangrijk, ga op zoek naar de bouwers, bedenkers, bezoekers en de vroegere gebruikers van het digitale object. Zij hebben de kennis en ervaring met het object!

TIP: Bouw een netwerk op

Breng je verhaal naar buiten en durf vragen te stellen aan experts in het veld (bel, mail, tweet, etc.). Dit heeft ons in het DDS-project naast kennis ook een geweldig netwerk en goodwill opgeleverd. Via het netwerk konden we ons verhaal op verschillende (internationale) podia vertellen, wat ons netwerk nog meer vergrootte en toegang bood tot nog meer kennis. De experts versterkten ons in onze missie en door het delen van hun kennis konden wij ons plan aanscherpen en verbeteren. Zonder deze kennis waren wij niet waar we nu zijn. Bereid je ook voor op nieuwe werkwijzen en samenwerkingsverbanden. Bestaande procedures en processen voldoen waarschijnlijk niet meer. Wees flexibel.

Zoek vanaf het begin de samenwerking met partners uit verschillende disciplines uit het digitale erfgoedveld, waaronder onder andere geesteswetenschappers, informatici, digital humanitiesspecialisten, conservatoren en software ontwikkelaars. Denk ook aan iemand die verstand heeft van digital forensics en/of E-discovery.⁵¹ Bekijk ook alvast het aanbod van de verschillende e-depots. Welke dienstverlening en infrastructurele voorzieningen heb je straks nodig?

Onze aanpak: case study 'DDS Herleeft'

'De Digitale Stad Herleeft' is ontwikkeld door Amsterdam Museum, Waag Society, het Nederlands Instituut voor Beeld en Geluid en de Universiteit van Amsterdam. De samenwerking van de vier projectpartners levert een bijzondere combinatie van kennis en expertise op, waarbinnen erfgoed, creatieve industrie, innovatie en technologie samenkomen. De werkgroep van de historicus van de wiskunde en informatica Gerard Alberts aan de Universiteit van Amsterdam is specialist op het gebied van geschiedenis van informatica, het Nederlands Instituut voor Beeld en Geluid is leidend in Nederland als het gaat om AV-archivering en sinds kort ook webarchivering, het Amsterdam Museum is de aangewezen partij om dit belangrijke deel van de Amsterdamse geschiedenis te presenteren aan het publiek en is specialist op het gebied van collectiebeheer en collectie-ontsluiting en Waag Society is nauw verbonden met de geschiedenis van DDS en is expert in het leiden van technische projecten waarbinnen verschillende disciplines elkaar vinden.

Naast projectpartners zochten we de oorspronkelijke inwoners van de stad op: de bedenkers, bouwers, bewoners en bezoekers van DDS. Zonder hun kennis, ervaringen en donaties hadden we geen basis gehad voor onze opgravingen en reconstructies. De herinneringen en verhalen van oud-bewoners en oud-medewerkers zijn de handleidingen voor de toekomst. Crowdsourcing is een onmisbaar instrument in webarcheologie. Naast de projectpartners en oorspronkelijke bewoners konden we bogen op een aantal enthousiaste vrijwilligers die ons op velerlei vlak hebben geholpen.

⁵¹ Er bestaat een grote tak binnen de IT met de naam 'digital forensics'. Een zusje hiervan heet e-discovery. Digital forensics wordt voornamelijk door LEA (Law Enforcement Agencies) ingezet om datadragers van onder andere verdachten uit te lezen en te analyseren. Daarnaast wordt het natuurlijk ook gebruikt voor data-recovery. E-Discovery heeft zijn oorsprong in de Verenigde Staten. In Noord-Amerika heb je bij rechtszaken tussen private partijen de zogeheten 'discovery phase'. Dit houdt in dat, in bepaalde gevallen, advocaten van een eisende partij aan de rechter kunnen vragen alle documenten te mogen kopiëren die relevant zijn voor de rechtszaak. Dit ging vroeger analoog, maar tegenwoordig gaat dit ook digitaal, hence de 'e' in e-discovery.

‘De Archeologische Dienst’

De speurtocht naar de partners werd vergemakkelijkt doordat wij (Tjarda de Haan en Paul Vogel) vroeger bij De Digitale Stad gewerkt hadden. We nemen je terug in de tijd: In 1997 verrichtten de webarcheologen van het allereerste uur en toenmalig DDS-werknemers, Jeroen van Kan en Nina Meilof, al de eerste opgravingen voor De Archeologische Dienst.⁵² Zeven jaar later, in 2004, pakten ze voor het tienjarige bestaan van DDS, de traditie weer op met de voortzetting van de Archeologische Dienst.⁵³ Ze openden het allereerste virtueel openluchtmuseum met archeologische resten van DDS. In samenwerking met Marleen Stikker, directeur van Waag Society en de eerste virtuele burgemeester van DDS, openden ze een virtuele brievenbus om objecten te verzamelen voor het openluchtmuseum. In de brievenbus kon iedereen DDS-gerelateerd materiaal opsturen. Waag Society organiseerde daarnaast ook nog een bijeenkomst om het tienjarige bestaan te vieren (een terugkerend evenement, bij het vijfjarige bestaan in 1999 was Waag Society ook betrokken). Na deze bijeenkomst kwamen Jeroen van Kan, Nina Meilof, Paul Vogel en Tjarda de Haan bij elkaar om te brainstormen over een mogelijke nieuwe toekomst van De Archeologische Dienst. We organiseerden gesprekken met verschillende partners, maar het was toen nog te vroeg om tot actie over te gaan en de plannen belandden in een bureaula.

Zes jaar later, in 2010, was er meer ‘historische afstand’. Het Amsterdam Museum (voorheen het Amsterdams Historisch Museum) bood als ‘neutrale’ organisatie een veilige haven en als museum de mogelijkheid om het DDS als object op te nemen in de collectie en te laten zien aan het publiek. De volgende generatie Archeologische Dienst was geboren. Via via, door social media, mailinglijsten en de eerder betrokken partners werd een eerste contactlijst van betrokken oud-werknemers samengesteld, alsmede een eerste inventarisatie van wat waar te vinden zou moeten zijn.

⁵² De Archeologische Dienst (1.0) verrichtte in 1997 al de eerste opgravingen, zie: <http://web.archive.org/web/19970615203717/www.dds.nl/dds/archeo.html>.

⁵³ De Archeologische Dienst (2.0) opende tijdens het 10-jarige bestaan het allereerste virtueel openluchtmuseum, zie: <http://www.vankan.dds.nl/dds/verjaardag.html>. Lees ook: De Archeologische Dienst 3.0 <https://hart.amsterdam/nl/page/33527/de-archeologische-dienst-3-0>

5.5 Maak een plan van aanpak

Nadat je de doelstellingen en de kennisgebieden in kaart hebt gebracht maak je een plan van aanpak. Dit is de roadmap voor jouw project.

Onze aanpak: case study 'DDS Herleeft'

De onderdelen in ons plan van aanpak zijn in de loop van het project aangescherpt door voortschrijdend inzicht maar zijn in hoofdlijnen hetzelfde gebleven.

De opzet van ons plan van aanpak zag er zo uit:

1. Start:
 - a. Schrijf een plan van aanpak
 - b. Zet een communicatieplatform op
2. Graven:
 - a. Verzamel mensen, verhalen en objecten
 - i. Crowdsourcing 1: Opening virtueel museum
 - ii. Crowdsourcing 2: De Grave Diggers Party
3. Reconstructie
 - a. Doe vooronderzoek
 - b. Doe onderzoek en reconstrueer
 - i. Maak data toegankelijk
 - ii. Verken de data
 - iii. Kies je doel
4. Oplevering en presentatie:
 - a. Archiveer DDS in e-depot(s)
 - b. Let the bytes free en maak DDS toegankelijk voor publiek!
5. Conclusies:
 - a. Conclusies, evaluatie, documentatie, kennis delen

5.6 Stel een begroting op

Na het opstellen van het plan van aanpak ga je verder met de begroting. De volgende stap is het vinden van bereidwillige fondsen. DEN, het Kennisinstituut Digitale Cultuur heeft een handige subsidiewijzer⁵⁴ opgesteld waar je tips vindt over het hoe en waarom van subsidieaanvragen en een overzicht van fondsen en subsidieregelingen voor digitaal erfgoed. Waag ook een stapje buiten de deur. Ook de Europese Unie biedt interessante regelingen, kijk bijvoorbeeld bij het 'Joint Programming Initiative in Cultural Heritage & Global Change' en 'Horizon 2020, The EU Framework Programme for Research and Innovation, die zich beiden ook richten op het behoud van cultureel erfgoed'.⁵⁵

⁵⁴ <http://www.den.nl/pagina/226/subsidiewijzer/>

⁵⁵ <http://www.jpi-culturalheritage.eu/> en <https://ec.europa.eu/programmes/horizon2020/>

Onze aanpak: case study 'DDS Herleeft'

Na volharding en met de juiste partners vonden wij het Prins Bernhard Cultuurfonds, Stimuleringsfonds Creatieve Industrie, het Netwerk Digitaal Erfgoed en het Mondriaan Fonds bereid om ons financieel te ondersteunen. Zij hebben verschillende regelingen voor e-cultuurprojecten.⁵⁶

5.7 Zet een communicatieplatform op

Maak een website om iedereen op de hoogte te brengen en houden van je project en de ontwikkelingen.⁵⁷ De site die je maakt kan een aparte site zijn onder een eigen domeinnaam of de site kan een onderdeel zijn van de website van je organisatie. Deel de content met alle projectpartners zodat je een groot bereik hebt. Zorg voor voldoende updates zodat het bestaan van het project tijd krijgt om wortel te schieten. Door je stappen te delen op een blog kunnen mensen het project volgen en komt het project tot leven. Een enkel persbericht is snel vergeten. Vergeet niet updates ook te crossposten op sociale media. Houd een lijst bij met de social media en mailinglijsten via welke je geïnteresseerden op de hoogte houdt van je project. Op deze manier ga jij naar het publiek toe (en wacht je niet tot ze naar jou toekomen) en geef je kattebelletjes zodat mensen weten dat er nieuws is en ze reden hebben je blog/website weer te bezoeken. Maak een contactlijst en onderhoudt deze gedurende het project. Verstuur de uitnodigingen per mail met een duidelijke afzender zodat mensen contact met je kunnen opnemen voor vragen, tips en suggesties.

Onze aanpak: case study 'DDS Herleeft'

Om iedereen op de hoogte te houden van de ontwikkelingen, opgravingen en bijeenkomsten ontwikkelden we een website, onder de naam re:DDS, de REconstructie van De Digitale Stad.⁵⁸ We richtten de site in met een blog, een Historisch depot voor de archeologische vondsten, een agenda, een tijdlijn om de geschiedenis van DDS in kaart te brengen, de rubriek 'WHOIS', een plek om herinneringen van de mensen achter de eerste virtuele stad van de wereld te verzamelen, de bouwers, bedenkers, bezoekers en bewoners) en met een uitleg van het project in 'Over het project'. De social media die wij inzetten zijn Twitter, Facebook, Youtube en Flickr.

⁵⁵ <http://www.jpi-culturalheritage.eu/> en <https://ec.europa.eu/programmes/horizon2020/>

⁵⁶ Prins Bernhard Cultuurfonds: <http://www.cultuurfonds.nl/aanvragen>, Stimuleringsfonds Creatieve Industrie: https://stimuleringsfonds.nl/en/grants/grant_programme_for_digital_culture/, het Netwerk Digitaal Erfgoed: <http://www.den.nl/nde> en het Mondriaan Fonds: <https://www.mondriaanfonds.nl/aanvragen/>.

⁵⁷ Dit communicatieplatform is onderdeel van je communicatieplan en dit valt buiten de scope van dit DIY.

⁵⁸ <https://hart.amsterdam/re-dds>

Onze aanpak: case study 'DDS Herleeft'

Om iedereen op de hoogte te houden van de ontwikkelingen, opgravingen en bijeenkomsten ontwikkelden we een website, onder de naam re:DDS, de REconstructie van De Digitale Stad. We richtten de site in met een blog, een Historisch depot voor de archeologische vondsten, een agenda, een tijdlijn om de geschiedenis van DDS in kaart te brengen, de rubriek 'WHOIS', een plek om herinneringen van de mensen achter de eerste virtuele stad van de wereld te verzamelen, de bouwers, bedenkers, bezoekers en bewoners) en met een uitleg van het project in 'Over het project'. De social media die wij inzetten zijn Twitter, Facebook, Youtube en Flickr.⁵⁹

TIP: Zet een duurzaam dossier op

Wellicht wordt jouw project net als de onze ook een 'slow data project' (we zijn inmiddels net zo lang aan het opgraven als dat DDS online is geweest). Daarom is het belangrijk dat je een goed, betrouwbaar, volledig en up-to-date dossier (werkomgeving) opzet met alle verzamelde informatie gedurende het project. Een belangrijk onderdeel van het dossier is de documentatie. Zorg voor goede documentatie. Leg alle stappen van je project zorgvuldig vast. Maak, indien mogelijk, ook foto's of video's. Belangrijk is dat je informatie snel en makkelijk kunt terugvinden, zodat iedereen binnen je organisatie (of project) het dossier eenvoudig kan gebruiken en aanvullen. Bijkomstig voordeel is dat je het eventueel kunt overdragen aan een opvolger en op termijn wellicht zelfs aan een archiefinstelling. Nog een aantal adviezen. Zet een overzichtelijke mappenstructuur op. De mappen geven context aan de bestanden die eronder zijn geordend. Orden niet op documenttype, maar op proces, onderwerp, en/of onderdelen van het project.⁶⁰ Gebruik betekenisvolle bestandsnamen. Verwerk hierin: type document, beschrijving, auteur, datum en status.⁶¹ De grootste dreiging voor je digitale archief is dat je na een tijd niet meer weet wat de juiste of de laatste versie van een document is. Voorkom dat er straks vele digitale kopieën op het internet of in je archief zwerven. Bewaar je bestanden in duurzame bestandsformaten. De vuistregel is hier: werk consequent.

TIP: Maak het project duurzaam

Creëer handvaten in de organisatie voor vergelijkbare toekomstige projecten. Schrijf handleidingen van bovengenoemde werkwijzen tijdens het project en plaats ze op een plek waar iedereen bij kan, bijvoorbeeld in het duurzaam dossier, en zorg dat iedereen zich hieraan houdt. Investeer in de kennisontwikkeling van de medewerkers en houd de ontwikkelingen bij door het volgen van blogs, seminars en participatie in (inter)nationale expert communities en deel deze.

⁵⁹ <https://twitter.com/amsterdammuseum>, <https://twitter.com/latjarda>, <https://www.flickr.com/photos/ahmamsterdam/albums/> en <https://www.youtube.com/user/reDeDigitaleStad>

⁶⁰ Bekijk de "Handleiding digitaal archiveren voor particulieren. Het eenvoudig beheren en (her)gebruiken van uw informatie met behulp van een mappenstructuur" in het dossier 'Digitale Archivering voor particulieren' op: <https://www.amsterdam.nl/stadsarchief/organisatie/archiefzorg-0/digitale-archivering/>

⁶¹ Karin van der Heiden ontwikkelde in samenwerking met Premsele, Nederlands Instituut voor Design en Mode, een handleiding met vuistregels voor het duurzaam archiveren van digitale bestanden. Bekijk: <http://bewaarals.nl/>

06

Graven

Hoe vind je verloren gegaan digitaal erfgoed terug?
Hoe graaf je digitale objecten op?

In deze fase ga je op zoek naar je digitale object, de artefacten en de overblijfselen. Dit is een van de spannendste onderdelen. Je gaat de grond in en maakt je handen vies. Hier kom je op het spoor wat er daadwerkelijk nog bestaat, en val je van de ene verrassing in de andere. Soms wekken vondsten helaas hoge verwachtingen die later in een teleurstelling eindigen. Blijf dus nuchter. Het graafproces kan chaotisch verlopen. Het is daarom van groot belang om tijdens de graafwerkzaamheden alle stappen gedisciplineerd te documenteren en ervoor te zorgen dat je je vondsten veilig stelt.

Stappenplan:

- 1 Maak een inventarisatie van gewenste objecten
- 2 Maak je plan wereldkundig
- 3 Maak vondstkaarten
- 4 Breng mensen, verhalen en objecten bij elkaar
- 5 Garandeer de data-integriteit van je vondsten
- 6 Stel je vondsten veilig

In de volgende paragrafen stappen we hier en daar af van het stappenplan en vertellen we eerst hoe wij het gedaan hebben aan de hand van onze case study.

6.1 Maak een inventarisatie van gewenste objecten

Je hebt vooronderzoek gedaan en je (digitale) object en het belang ervan beschreven. De volgende stap is een wensenlijst opstellen met objecten die je wilt vinden.

Onze aanpak: case study 'DDS Herleeft'

Nadat we de eerste stappen in het terugvinden van de oorspronkelijke bewoners en de overblijfselen van de verloren stad gezet hadden, stelden we een eerste wensenlijst op. We hoopten de volgende archeologische resten te vinden: de drie DDS interfaces (DDS 1.0,2.0,3.0), de servers (Shaman, Alibaba, Aladdin en Sarah), de tape robot, De FREEZE uit 1996 (hierover later meer), de usenet nieuwsgroepen (oa dds.dds, dds.multicult, dds.technopolis), discussiefora, de Metro, de cafés, DDS Webmix, live.dds.nl, Digitale Stedeling, Digitale Huiskamer etc. Tijd om te gaan graven en de kracht van crowdsourcing in te zetten!

6.2 Maak je plan wereldkundig

Organiseer een kick-off bijeenkomst en nodig iedereen uit die zich verbonden voelt met je project. Hiermee stimuleer je participatie, communityvorming, samenwerking, interactie, crowdsourcing en je vergroot je netwerk. Deel je stappen via je communicatieplatform.

Onze aanpak: case study 'DDS Herleeft'

Op 8 april 2011 was het inmiddels tien jaar geleden dat het internet-pioniersproject en de oudste virtuele gemeenschap in Nederland, De Digitale Stad, uit het publieke domein verdween. Hoog tijd, en een mooie mijlpaal, om te gaan graven. We nodigden iedereen, die we inmiddels hadden opgespoord, uit voor de kick-off van ons project en de lancering van ons virtuele museum. De kick-off vond plaats op de legendarische Prins Hendrikkade 194 in Amsterdam. Dit is de plek waar DDS zich gevestigd had na de start in De Balie: Bij uitstek de plek om oud-DDS-'ers te herenigen op een reünie en te vertellen van ons plan. We nodigden iedereen uit om een drankje te komen drinken en om mee te doen: Word ook digitaal archeoloog en help de DDS op te graven. Een sociaal evenement (met bier, dat brengt de mensen samen en levert mooie anekdotes op! En ja, appelsap werkt ook!) met vroege en toekomstige stakeholders van je project werkt zeer goed. Iedereen is op de hoogte van het plan en gaat nadenken over mogelijke bijdragen. Door het samenbrengen van de bouwers, bedenkers, bezoekers en de gebruikers van DDS ontstaat kennisoverdracht en een beter begrip. Samen ben je slimmer. De kracht van crowdsourcing! Je komt er ook achter

waar artefacten eventueel zijn. En dan kun je mensen benaderen. We kondigden ook meteen ons tweede evenement aan, een maand later: de ‘Grave Diggers Party’. Hierover straks meer.

Door de -inmiddels binnendruppelende- teruggevonden artefacten online te plaatsen in het virtueel openluchtmuseum kwamen ze weer tot leven. Sommige gevonden voorwerpen waren verloren gewaand en werden via donaties weer boven water gehaald. Unieke stukken zijn zo weer terug gevonden, en dienden weer als aanwijzingen voor nieuw te zoeken onderdelen.

Tijdens de kick-off hadden we iedereen uitgenodigd om de rubriek ‘WHOIS’ in te vullen. De naam komt van de ‘Whois’, de tool die we al eerder beschreven. In de ‘WHOIS’ (hiervoor gebruiken we een Google Doc formulier) stellen wij een aantal vragen aan iedereen die herinneringen heeft aan DDS. Bijvoorbeeld: Help ons herinneren. Wat moeten we opgraven? Wat betekende DDS voor jou? Wat heb jij nog thuis en wil je doneren? En meer. Op de vraag “Wat zou jij graag in een (fysieke) tentoonstelling terugzien?” kregen we onder andere de volgende antwoorden: “De terminals, de servers, foto's van de eerste bewoners, hun avatars, de interfaces, de ideeën die geopperd werden. DDS draaiend op Netscape op pc's uit die tijd. Alle bytes die ik verloren ben. Een reconstructie van de site van destijds... de digitale Metro. De oude site. Als hommage. De mogelijkheid om te chatten zoals toen. Grote platen van de DDS interface. Werkende DDS. Eigenlijk de sfeer. Maar dat wordt lastig. The Stone Age Computer. DDS 3.0 werkend. De Bruine Kroeg.” Mooie aanwijzingen om verder te zoeken. Over de kick-off van het project en de opening van ons virtueel museum hebben we regelmatig gecommuniceerd op onze website en via social media. Vuistregel is: Deel, waar en zo vaak mogelijk, je kennis!

TIP: Bezoek eens een reünie

Heb je te maken met een groep mensen (webarcheologen, IT'ers, providers of erfgoedprofessionals of -liefhebbers) die soms een reünie organiseren? Sluit je dan daarbij aan, aangezien je dan meteen de kanalen hebt en adressenlijsten kunt samenstellen. Je kunt geïnteresseerden voor jouw bijeenkomsten uitnodigen. Opkomst gegarandeerd!

6.3 Maak vondstkaarten

Opgraven lijkt een soort schatgraven, maar er komt ook flink wat werk bij kijken. Het beschrijven van je vondsten bijvoorbeeld. Een museum registreert en documenteert elk object in de collectie. Je wilt alle informatie over het object verzamelen, bijvoorbeeld: Wat is het, hoe werd het gebruikt, hoe is het verkregen, van wie is het, zijn er bijzonderheden (voor het beheer bijvoorbeeld)? Vondstkaarten dienen als basis voor de documentatie. Je gebruikt deze informatie ook voor onder andere doeleinden, zoals acquisitie, onderzoek en overdracht naar een collectiedatabase van het museum en het e-depot. Als je voor een museum werkt, dan gebruik je ongetwijfeld al vondstkaarten die als basis kunnen dienen.

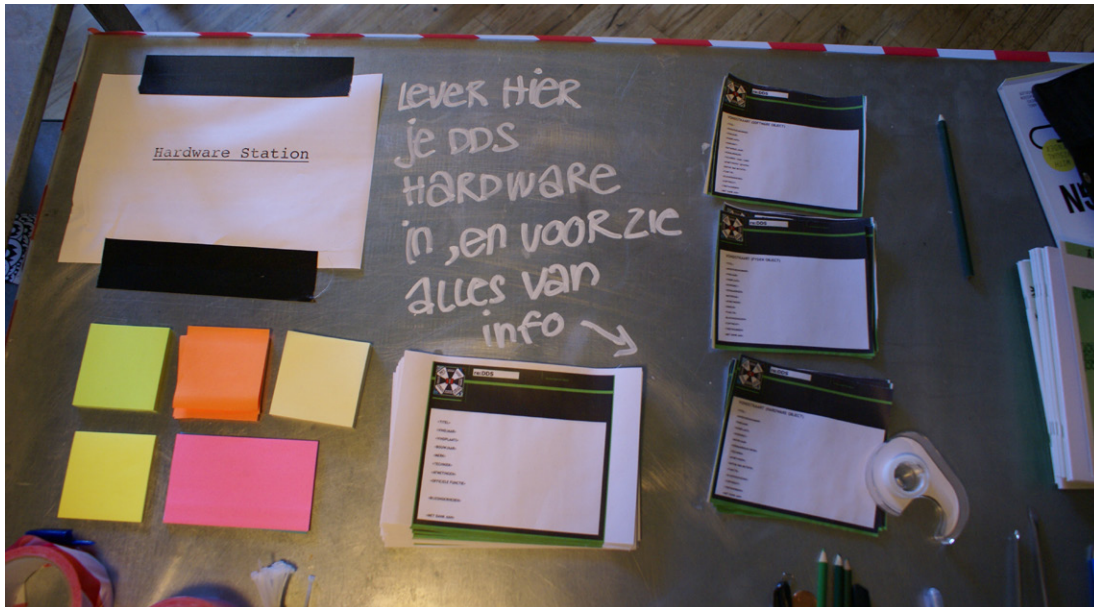
Onze aanpak: case study 'DDS Herleeft'

Door de eerste bijeenkomst en oproepen ontvingen we de eerste donaties. Om de vondsten te beschrijven maakten we vondstkaarten voor fysieke-, software- en hardware 'opgravingen'. De vondstkaarten plaatsen we in ons online 'Historisch depot', met waar mogelijk inclusief de gedigitaliseerde oorspronkelijke documenten. Zo is een deel van de geschiedenis van het publieke domein deel van DDS weer toegankelijk. Door het OCR scannen van de documenten zijn ze (weer) vindbaar en doorzoekbaar.

Per opgraving beschrijven we:

- Titel
- Vindjaar
- Vindplaats
- Herkomst
- Jaar van vervaardiging (bouwjaar bij hardware)
- Vervaardiger (vervaardiger/merk bij hardware)
- Materialen (techniek bij hardware en formaat bij software; en versie bij beiden)
- Afmetingen
- Inhoud
- Functie (de werking van het object en het gebruik)
- Bijzonderheden
- Copyrights
- Met dank aan

Bij de vondstkaart van Hardware voegden we nog 'Datum van mutatie' (wijziging) toe. En onder elke vondstkaart die we online zetten, plaatsten we de oproep: "Wie weet meer over deze vondst? Help ons en plaats een reply met jouw bevindingen of stuur een mail!" Dit heeft ons veel informatie opgeleverd. Nogmaals, crowdsourcing is een essentieel onderdeel bij het opgraven van de beginjaren van het web.



Vondstkaarten

6.4 Breng mensen, verhalen en objecten bij elkaar

Inmiddels heb je bronnenonderzoek gedaan, heb je een wensenlijst met objecten en heb je een contactlijst met deelnemers. Het is nu tijd om de kracht in van crowdsourcing in te zetten. Vind een locatie waar je de graafwerkzaamheden kunt uitvoeren en stel een team samen met technici, historici, bloggers/vloggers en vrijwilligers als vliegende keep. Tijd voor een feestje!

Onze aanpak: case study 'DDS Herleeft'

Op vrijdag 13 mei 2011 organiseerde het Amsterdam Museum samen met Waag Society de 'Grave Diggers Party' in Amsterdam. Een feest om mensen te herenigen en verloren herinneringen, zowel persoonlijk als digitaal, te verzamelen. Het samenbrengen van alle mensen, met verschillende expertises, was op deze bijeenkomst zeer belangrijk en noodzakelijk. Het digitaal erfgoed, en zeker het complexe digitale object dat DDS is, vergt veel (vooral technische) ervaring en expertise. Niet alleen hedendaagse kennis, juist de kennis van de historische data, software en hardware is essentieel. Een bijeenkomst is ook heel praktisch: Doordat je alle mensen op één datum op één plek uitnodigt voorkom je dat je bij iedereen apart langs moet gaan. De locatie Waag Society op de Nieuwmarkt in Amsterdam was makkelijk bereikbaar voor iedereen en als medialab uitgerust met alle benodigdheden. Ook de interactie tussen de aanwezigen zorgt ervoor dat oude herinneringen bovenkomen.

We nodigden dus alle voormalige inwoners, oud-DDS werknemers en zielsverwanten uit voor de 'Grave Diggers Party': "Help ons om deze unieke stad op te graven, kom naar de aftrap van onze webarcheologische

graafwerkzaamheden en draag bij aan de restauratie van dit historisch monument!". "Kijk op je zolder en/of harde schijven. Breng alle servers, modembanken, VT100-terminals, freezes, harde schijven, scripts, zips, floppies, tapes, back-ups, logbestanden, video's, foto's en screenshots die je kunt vinden en vertel ons je herinneringen en verhalen!" Vijftig enthousiaste (sommige internationale) cybernauten reageerden op onze oproep en kwamen langs met volle tassen, harde schijven en USB-sticks.

We hadden een 'Archeologische Vindplaats' opgezet in het Zuidlab van Waag Society waar de digitale opgravingen werden gedaan. Daarnaast hadden we een 'Tijdelijke Tentoonstelling' opgebouwd waar we de gevonden artefacten verzamelden en aan het publiek toonden. Onze 'Archeologische Vindplaats' bestond uit twee werkruimtes. Eén voor de sociale webarcheologie en één voor de technische webarcheologie. In de sociale vindplaats lieten we deelnemers op zoek gaan naar hun herinneringen op het web om een duidelijker beeld te krijgen van de toenmalige DDS (de samenleving, de sociale context) en hoe ze DDS (de media en informatie: de stad met pleinen, huizen, zijwegen, magazine, cafés en discussies) gebruikten. We plaatsen werkstations met de Wayback Machine naar de DDS website en de huizen. Terwijl deelnemers door de historische stad en huizen surfden, door de tijd heen, bedreven we 'oral history' om verhalen te verzamelen. Oral history is een vorm van geschiedschrijving om herinneringen van ooggetuigen vast te leggen. Dit deden we aan de hand van interviews op basis van de WHOIS vragen. Het veldonderzoek leverde ons waardevolle bijdragen voor onze wensenlijst op om op te graven.

TIP: Ga aan de slag met 'oral history'

Interview de betrokkenen ter plekke. Dit levert mooie herinneringen op, en zorgt ervoor dat de gesprekken loskomen. Deze verhalen, met de kennis en ervaringen uit het verleden, zijn de handleidingen voor de toekomst. Het geheugen waarover de deelnemers beschikken staan doorgaans niet in boekjes en kunnen enorm waardevol zijn voor je project. Maak eventueel afspraken om deelnemers later te interviewen. Vergeet niet een videocamera mee te brengen voor de interviews. Leg alle stappen zorgvuldig vast en fotografeer de graafwerkzaamheden, de vindplaatsen en de vondsten.

De tweede werkruimte bestond uit de technische vindplaats en zag eruit als een crime scene. Midden in het Zuidlab stonden tafels met werkstations afgezet met rood-wit lint. Dit was het exclusieve werkgebied van de technische webarcheologen. Gewapend met digitaal gereedschap spitten zij in het aangeleverde historisch DDS-materiaal om de geschiedenis en het gebruik van vroege computerarchitectuur, randapparatuur, besturingssystemen, programmeertalen, systeembeheer tools en dergelijke te onderzoeken. Het gereedschap bestond uit computers (digitale graafmachines), opslagapparatuur (emmers), USB-sticks (vondstzakken), UNIX-commando's (spades, pikhouwelen, troffels) en scripts (metaaldetectoren). De webarcheologen verzamelden de gevonden voorwerpen, de (oorspronkelijke) overblijfselen van de verloren gegane virtuele stad, op een harde schijf.



'Archeologische Vindplaats' op de 'Grave Diggers Party'

TIP: Laat deelnemers de machine (computer) meenemen

Data worden hoogstwaarschijnlijk meegenomen naar de bijeenkomst op verschillende datadragers, bijvoorbeeld in de vorm van een USB-stick, harde schijven of zelfs tapes. Laat iedereen, als dit mogelijk is, altijd de gehele machine (computer) meenemen waarop de vondsten staan, zodat de technische webarcheologen ervoor kunnen zorgen dat alle data goed overgezet worden. Waarom dit belangrijk is leggen we uit in de volgende paragraaf 'Stel je vondsten veilig'.

Technische verdieping: Maak een lijst met graafgereedschap⁶²

Dit was de kit-list van de techneut die technische bijstand verleende bij de 'Grave Diggers Party'. Deze tools werden gebruikt om data van oude SUN systemen af te halen:

- Notebook met internetverbinding (voor het opzoeken van technische documentatie)
- Desktop met verschillende SCSI-kaarten en Knoppix live DVD (voor het overzetten data)
- Install disc's met Solaris 10 en oudere versies (SunOS 2.4 en nieuwer)
- SUN-keyboard en monitor
- Connector: PC serial naar RS232 (nullmodem kabel)
- Gender benders RS232
- RS232 9 pin (DE-9) male naar 25 pin (DB25) female kabel
- SCSI (Small Computer System Interface) converter
- SCSI 1 naar Scsi 2 converter (en elke andere SCSI connector mogelijk)
- SCSI 2 kaarten
- Switches (ethernet; moet backwards compatible zijn met oude ethernet snelheden)
- USB 2 serial converter
- (USB) storage: minimaal 2 keer 1 terabyte
- UTP kabels
- Stekkerdozen en extra server (power) kabels
- Gereedschapskist met schroevendraaiers
- busje perslucht (geen overbodige luxe)

TIP: Verzamel je webarcheologische gereedschapskist op tijd

Sommige (historische) onderdelen zijn niet makkelijk meer te verkrijgen. Wellicht moet je uitwijken naar online veiling-websites, zoals eBay of Marktplaats. Of je kunt natuurlijk een oproep doen via je communityplatform en/of je social media kanalen.

⁶² Twee handige websites: <http://classiccomputers.info/manuals.html> bevat oude technische handleidingen. <https://amiaopensource.github.io/cable-bible/> is een handleiding voor kabels en aansluitingen die gebruikt worden voor het behoud van het audiovisueel erfgoed. De Cable Bible is een project van het AMIA Open Source Committee.

TIP: Houd de kabels in de gaten

Bij oude hardware is het cruciaal om bij te houden welke kabel waarin of waarbij hoort. Gebruik labels en foto's om dit bij te houden. In ons geval hadden we te maken met externe SCSI-packs die alleen in een bepaalde volgorde werkten. Let op veroudering van kabels en plastic.

Na de 'Grave Diggers Party' diende Waag Society nog drie weken lang als interactieve archeologische vindplaats. Iedere vrijdag konden deelnemers langskomen om veldwerk te verrichten of van een rondleiding te genieten in de 'Tijdelijke Tentoonstelling'. En de tentoonstelling groeide en groeide. Tijdens de toeristische bezichtigingen interviewden we de deelnemers die door de gevonden voorwerpen weer verhalen herinnerden.

Gevonden voorwerpen

Na drie weken graven hadden we zeer bijzondere artefacten gevonden. Waag Society vond en schonk twee publieke terminals. De terminals zijn ontworpen door Studio Stallinga in 1994. De terminals waren publieke pronkstukken van de virtuele stad. Voor wie thuis geen computer met een modem had, kon gratis gebruik maken van de terminals op verschillende openbare plekken in Amsterdam. De terminals stonden in het stadhuis, het café van De Balie, het Stedelijk Museum, de Openbare Bibliotheek, het AMC en het Amsterdams Historisch Museum. Volgens verhalen op de 'Grave Diggers Party' waren de terminals zo populair dat er rijen mensen stonden te wachten om plaats te kunnen nemen in de ijzeren stoel om rond te kunnen dwalen in De Digitale Stad. DDS-programmeurs bouwden al snel een tijdslimiet in per persoon. Na 30 minuten werd je uit de stad gezet en was de volgende persoon aan de beurt.

Voormalige systeembeheerders van DDS brachten afgedankte servers mee die ze hadden gered uit vuilniscontainers. De servers droegen exotische namen zoals Alibaba, Shaman, Sarah en Alladin. Webarcheoloog Paul Vogel vertelde ons dat sommige DDS servers vernoemd waren naar de gerechten van de snackbar om de hoek van de DDS (naast het café Scharrebier, de stamkroeg van DDS en de lokale hacker community). Sommige servers verkeerden in een erbarmelijke staat. Gekannibaliseerd, dat wil zeggen leeggeroofd (de kosten waren zo hoog in de jaren negentig dat alles wat nog werkte werd hergebruikt), simpelweg kapot of gewist. Soms waren de oorspronkelijke data overschreven. Een aantal harde schijven was wel te lezen en bevatten nog data. We werkten met connectors die het mogelijk maakten om de data op moderne hardware uit te lezen. Qua setup gebruikten we een Linux-liveCD op een desktop, aangezien we daar de specifieke hardware mee konden uitlezen. Een notebook kan namelijk geen PCI-kaarten gebruiken, en heeft ook geen seriële poort, etc.

Een voormalige bewoner schonk een van de eerste modem-banken die in de eerste jaren van DDS dienst deed. Een andere bewoner stuurde een speciaal voor DDS gemaakt hoorspel uit 1994: 'Station Het Oor'. Het hoorspel werd tijdens de eerste zes weken na de opening van de stad uitgezonden. Het was gebaseerd op discussies en bijdragen van de eerste digitale stedelingen.

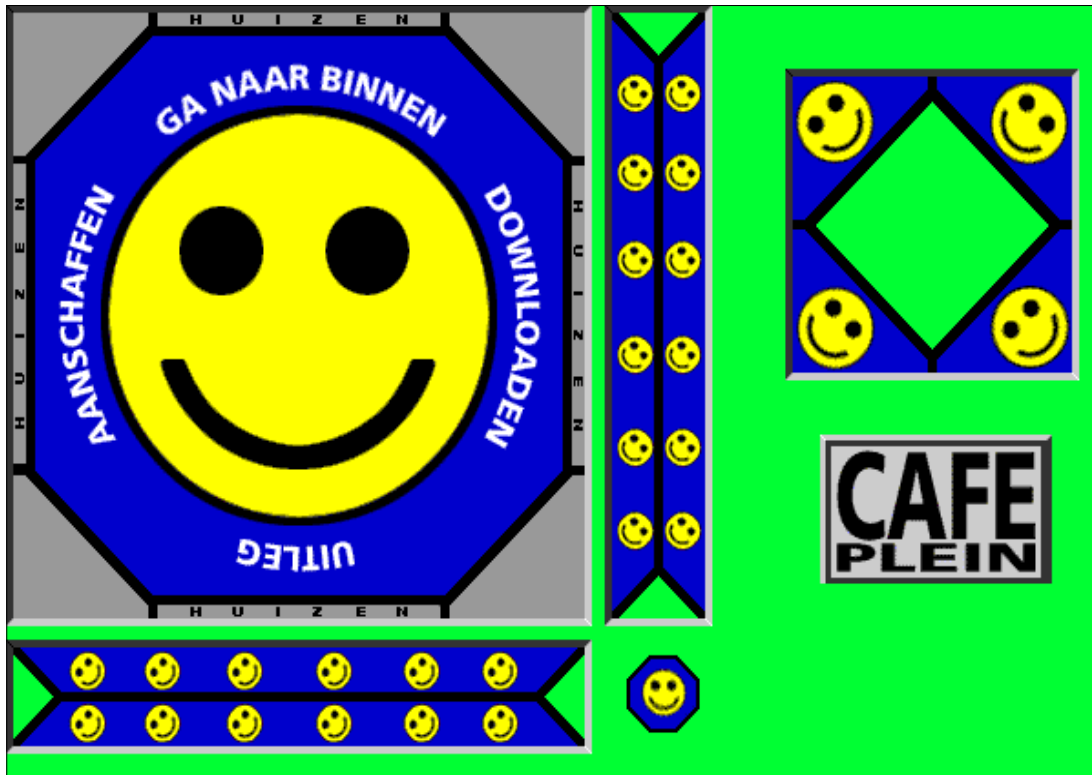


Ansichtkaart uit DDS, vondst op de 'Grave Diggers Party'

Dankzij de gulle donaties van oud-bewoners en oud-medewerkers hebben we inmiddels een rijke collectie DDS-artefacten. De archeologische resten van de stad bestaan uit:⁶³

- Gedigitaliseerd erfgoed: foto's, handleidingen, uitgave van de eerste Digitale Stedeling, documenten van de Gemeente Amsterdam, ansichtkaarten uit DDS, originele DDS-muismat, audiovisueel materiaal (van videobanden) waaronder vier afleveringen SMART TV, drie DDS Webmix afleveringen over De Archeologische Dienst en het internet hoorspel "Station Het Oor".
- Digitale informatie over erfgoed: vondstkaarten en metadata.
- Born digital erfgoed: de gedoneerde data tijdens de Grave Diggers Party op 13 mei 2011 (losse data/back-ups en de data van de harde schijven uit de gedoneerde servers). Hier werd een paar weken later De FREEZE uit 1996 (data tot en met 15 januari 1996) aan toegevoegd, zoals je kunt lezen in de volgende paragraaf. De FREEZE bestaat uit: nieuwsgroepen, de Metro, cafés, DDS3.0 met de pleinen en individuele huizen (tot 15 januari 1996; de datum van De FREEZE).
- Analoge artefacten: publieke terminals, servers (Shaman, Alibaba, Aladdin, Sarah), disken (Shaman, Disk 6, Disk 10) en een modem-bank.

⁶³ Bekijk het Historisch Depot: <https://hart.amsterdam/nl/page/33435/historisch-depot-re-dds>



De bouwplaat voor het plein DDS Centraal, ontwerp voor DDS 3.0, webarcheologische vondst

De webarcheologen hebben op de 'Grave Diggers Party' in totaal dertig gigabyte aan data opgegraven uit de nog 'levende' harde schijven van de servers en uit de verschillende back-ups van DDS die we gedoneerd kregen. We hebben de vondstkaarten gebruikt om de herkomst van de teruggevonden artefacten (analoog of digitaal) te documenteren. Hierdoor weet je wie het materiaal heeft ingebracht, waar het vandaan kwam, hoe het origineel werd gebruikt, wat de huidige staat van de objecten is en wie je kunt benaderen voor meer informatie, mocht dat nodig zijn. Deze 'data over data', de metadata, spelen een belangrijke rol bij het veilige behoud van de data, waarover later meer.

De FREEZE

Na de 'Grave Diggers Party' kregen we nóg een unieke vondst in handen. Joost Flint, voormalig directeur van DDS, vond drie DLT-tapes, 'alibaba freeze', 'SHAMAN (FREEZ)' en 'dds freeze', en gaf ons deze in bruikleen voor onderzoek. Samen vormen de tapes De FREEZE, waar we naarstig op naar zoek waren. Op 15 januari 1996 bestond DDS precies twee jaar. Voor de tweede verjaardag toonden de DDS-pioniers (weer) hun vooruitziende blik. Ze 'bevroren' DDS en stuurden deze als flessenpost de toekomst in. Ze beschreven en beargumenteerden hun verjaardagscadeau als volgt:

"In de afgelopen twee jaar is De Digitale Stad vooral gericht geweest op de toekomst. Deze verjaardag is een goed moment om eens stil te staan bij wat er de afgelopen twee jaar gebeurd is in de stad. Spraakmakende discussies, grootscheepse stadsuitbreidingen, komende en gaande bewoners, vriendschappen, liefdes en hooglopende ruzies, de Digitale Stad heeft al een hele geschiedenis."

(...)



De FREEZE

"Drie versies van De Digitale Stad zijn er al geweest in twee jaar. Mensen komen en gaan. Trends komen op en verdwijnen weer. Huizen worden opgetuigd en verfraaid. Maar waar blijven alle digitale data? Wie weet over 5 jaar nog hoe DDS 3.0 eruit zag?"
 (...)

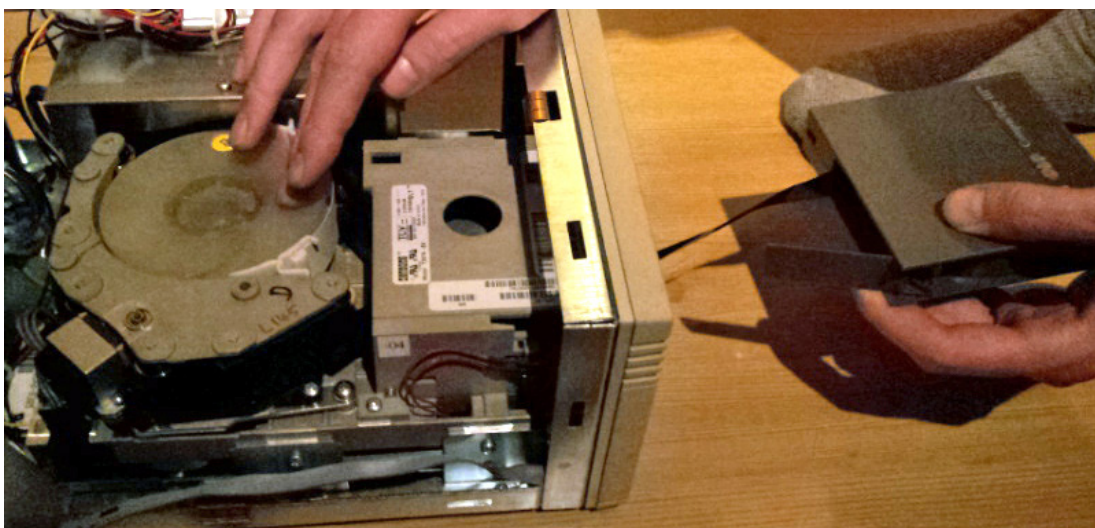
*"De Digitale Stad zal op maandag 15 januari om 18.00 precies 'bevroren' worden. Een momentopname, exact met alles wat de stad te bieden heeft zal tussen 18.00 en 9.00 de volgende ochtend worden opgeslagen op schijf en hermetisch worden verpakt. De schijven met data, samen met een complete beschrijving van de programma's en machines waarop de DDS nu draait zal worden gedeponeerd in **een archief ter bestudering voor archeologen in een verre toekomst.**"⁶⁴*

Na het enorme eureka!-gevoel van de unieke vondst kwam de ontnuchtering. Helaas voor ons waren (ondanks de beloften van de DDS pioniers die De FREEZE maakten), de documentatie van de programma's én de hardware waar de DDS toen op draaide niet bewaard. Maar na een verwoede speurtocht vonden we een 'antieke' tape reader om de drie tapes te kunnen uitlezen. Nadat we via e-mail en social media ons netwerk hadden bevraagd mailde de Koninklijke Bibliotheek dat ze een vondst hadden gedaan. Onder een bureau hadden ze de historische hardware gevonden die we zochten. De tape reader werd gebruikt als ... voetenbankje! "We zijn heel slecht in weggooien." De KB doneerde de reader voor het project.

⁶⁴ Bron: <http://web.archive.org/web/19991022030349/http://www.dds.nl/HyperNews/get.cgi/freeze.html>

In grote opwinding begonnen we met graven. Maar na twee nachten van 'reverse engineering' van de tape reader en de tapes, vonden we maar 70MB waar we hoopten minimaal 3GB te vinden. Teleurstelling alom. De geluiden die uit de tape reader kwamen (tok, prrrrr, beeeep) gaven ons de indruk dat de tapes waren gewist en we dachten dat de stad voorgoed verdampt was. We waren ontgoocheld. Zou er ooit meer gegevens op de tapes gestaan hebben? Is dit De FREEZE die we hadden gehoopt te vinden? Wat was stuk, de tapes reader of de tapes? Of allebei?

Na onze eigen mislukte pogingen de tapes uit te lezen en bloedstollende 'reverse engineering' ervaringen belden wij het Computermuseum van de Universiteit van Amsterdam voor hulp. De meeste machines in het museum werken nog en zijn uitgerust met de originele systeem- en applicatiesoftware en zijn volledig gedocumenteerd. Zou het museum in staat zijn om de historische data van De Digitale Stad van de tapes te lezen?



'Reverse engineering' van de tapes en de tape reader

```
805.765815] st0: Failed to read 8192 byte block with 1024 byte transfer
805.780518] st0: Failed to read 8192 byte block with 1024 byte transfer
805.795203] st0: Failed to read 8192 byte block with 1024 byte transfer
805.810547] st0: Failed to read 8192 byte block with 1024 byte transfer
805.825437] st0: Failed to read 8192 byte block with 1024 byte transfer
805.840605] st0: Failed to read 8192 byte block with 1024 byte transfer
805.855535] st0: Failed to read 8192 byte block with 1024 byte transfer
805.869900] st0: Failed to read 8192 byte block with 1024 byte transfer
805.884691] st0: Failed to read 8192 byte block with 1024 byte transfer
805.899480] st0: Failed to read 8192 byte block with 1024 byte transfer
805.913717] st0: Failed to read 8192 byte block with 1024 byte transfer
805.927838] st0: Failed to read 8192 byte block with 1024 byte transfer
805.942043] st0: Failed to read 8192 byte block with 1024 byte transfer
805.956493] st0: Failed to read 8192 byte block with 1024 byte transfer
```

#Fail

Henk Peek, de technicus van het Computermuseum, begon meteen met graven. Na een paar weken mailde hij ons een statusrapport: "Ik heb circa 11 Gigabyte data van de tapes gelezen. De tapes zijn 2 maal gelezen en de data daarna vergeleken. Er zijn geen data verschillen tussen de leeslagen. Het ziet er erg positief uit!"⁶⁵

⁶⁵ Lees: UvA Computermuseum graaft 11GB data op <https://hart.amsterdam/nl/page/37203/uva-computermuseum-graaft-11gb-data-op>

Technische verdieping: Bereid je voor op technische uitdagingen

In deze paragraaf beschrijven we een aantal problemen waar we gedurende ons project tijdens het recoveren van de data tegenaan zijn gelopen. Dat je een datadrager of een compleet systeem hebt, wil nog niet zeggen dat je de data ook kunt uitlezen. Dat je de data kunt uitlezen, wil nog niet zeggen dat je deze kunt herkennen. En dat je de data kunt herkennen, wil nog niet zeggen dat de data correct (intact) zijn.

Henk Peek van het Computermuseum was aan de slag gegaan met onze tapes. Hij koos ervoor om niet de tape reader van de KB te gebruiken, maar zette een eigen reader in. Deze heeft hij eerst op de 'hard way' moeten fixen. Vervolgens heeft hij alle tapes eerst handmatig schoongemaakt, voordat het -na enige moeite- lukte om de gevonden data uit te lezen. Tapes hebben de vervelende eigenschap dat het onmogelijk is te weten wat het einde is vanuit het perspectief van de reader. Als de tape reader dan ook 'lege data' tegenkomt, dan zal dat wel het einde van de tape zijn. Als er echter data verloren zijn, dan kan dit ook een gat geven, en dan probeert de tape reader niet eens verder te lezen. Ook in het tijdperk van de automatisering gaat niet alles automatisch. Het is aan de meer dan 40 jaar ervaring en het enorme doorzettingsvermogen van Henk Peek te danken dat we de data van de tapes hebben kunnen recoveren. Deze waren natuurlijk niet in een formaat dat vandaag de dag makkelijk gelezen kan worden, maar met de Solaris (SunOS) versie van restore was het binary formaat uit te lezen (de linux versie kon er niet meer overweg... SunOS was te oud).

Op de 'Grave Diggers Party', waar we oude systemen verzamelden en probeerden uit te lezen, ging ook niet alles van een leien dakje. De SCSI SCA-connectors bleken het zo goed als niet te doen (ze stopten na de eerst paar SCSI-requests), zodat we creatief moesten zijn. We konden de schijven uitlezen door het systeem onder stroom te houden zonder het OS te booten en de SCSI-bus aan te sluiten op de SCSI-bus van een ander systeem. Dit vereiste wel een paar hacks, omdat de SCSI-kaart standaard id 0 is. Daarnaast hadden we natuurlijk geen standaard disklabel support in de kernel (de partitie-standaard op SunOS-systemen), en de UFS-versie van het filesysteem kon niet gelezen worden door de Linux UFS-support. De systemen waren aangesloten via een seriële kabel (de Sun-monitor en keyboard was niet gelukt), maar veel systemen gaven geen enkel signaal. Was dit omdat de seriële port-instellingen niet correct waren? (Er zijn namelijk erg veel combinaties mogelijk), of was het systeem simpelweg niet meer werkend?

In een later stadium in het project probeerden we, inmiddels versterkt door een team van studenten, de harde schijven uit te lezen die weerbarstig alle pogingen om hun geheimen te ontsluiten hadden weerstaan. Na een oproep via de sociale media kanalen kregen we een oude werkende Sun-server gedoneerd. Deze was zeer onhandelbaar qua gewicht (loeizwaar), maar zorgde er wel voor dat de schijven die nog niet uitgelezen waren alsnog uitgelezen konden worden. We liepen wederom tegen data-corruptie aan. Zo had een bestand een incorrecte filesysteem attribuuft dat aangaf dat het bestand een paar honderd terabyte groot was, hetgeen het downloaden van deze data door de studenten (in de reconstructie fase) niet vergemakkelijkte.

Het werken met de data bleek lastig vanwege het ontbreken van een goede digitale werkomgeving met genoeg capaciteit (genoeg CPU's, genoeg geheugen, genoeg ruimte die snel genoeg is etc.). Een van de vrijwillige webarcheologen had een server in een datacenter met zeer veel vrije ruimte die ons tijdelijk van dienst was. De webarcheoloog doneerde later een andere server die 'toch thuis op de stapel niks lag te doen' aan het project. Deze werd uitgerust met nieuwe harde schijven (dankzij een klein budget), en opgehangen in het datacenter van een van de partners. Dit maakte het mogelijk voor het team om sneller en flexibeler te werken. Het opstarten van een formeel traject hiervoor zou betekend hebben dat er een groter budget en een langer traject nodig zou zijn geweest waarbij offertes opgevraagd hadden moeten worden, en na veel red-tape (en veel tijd) hetzelfde zou zijn bereikt. Dit had veel tijd gekost en mogelijke goodwill. Budget is overal eindig. Daarom ons advies: Zet het in op die plekken waar het echt nodig is.

TIP: Deel je vraagstukken en problemen met je community en je netwerk

Onderschat nooit het improviserend vermogen van je eigen mensen en je community. Deel daarom je vraagstukken en problemen snel en in een vroeg stadium. In het project heeft Waag Society in samenwerking met de andere projectpartners hiervoor een aantal Hacking Heritage Labs georganiseerd. Tijdens de labs hebben we samen met de community en het netwerk onderzocht hoe digitaal erfgoed gerestaureerd, bewaard en tentoongesteld kan worden. Hierdoor zijn we tot nieuwe kennis, oplossingen en inzichten gekomen die we direct konden toepassen.

6.5 Garandeer de data-integriteit van je vondsten

We maken nu eerst even een uitstapje naar de archeologie omdat een vergelijking heel verhelderend kan zijn. Archeologie is de wetenschap die overblijfselen van oude culturen bestudeert om het verleden te reconstrueren en te duiden. Stel, je haalt een oude kruik uit de bodem bij een opgraving. Je wilt dan niet alleen de historische kwetsbare kruik zo compleet mogelijk, en het liefst in één geheel, uit de bodem halen. Ook de bodem waarin de kruik zich bevindt herbergt veel informatie over de kruik. In de bodem kunnen bijvoorbeeld zaden zitten. De zaden zeggen iets over de planten en het klimaat waarin ze groeien. De sedimentlagen in de bodem kunnen weer iets zeggen over het grondwaterpeil. En met behulp van carbon-dating kunnen archeologen achterhalen hoe oud de aardlaag is waarin het artefact zich bevond. Alle informatie over de kruik én de context waarin de kruik is gevonden, zijn dus van belang.

Zo werkt het ook met metadata. Wees je ervan bewust dat er meer data zijn dan je in eerste instantie zou denken. Niet alleen het behoud van de data zelf is belangrijk, ook het behoud van de authenticiteit van de data is van groot belang. Om de herkomst van de digitale objecten zeker te stellen maak je gebruik van metadata. Er zijn verschillende metadata, en deze worden op verschillende momenten (tijdens het maken, veranderen en behouden van een digitaal object) toegevoegd. Bij het overzetten van de data is het dus essentieel om deze metadata te behouden.

Wat zijn nu precies metadata en welke rol spelen ze bij het veiligstellen van je digitale vondsten? Metadata zijn gegevens over gegevens, data over data. Waarbij het digitaal erfgoed de 'content' is, verschaffen de metadata de 'context'. Metadata helpen bij het begrijpbaar, vindbaar, bruikbaar, beheersbaar en authentiek houden van digitale bestanden. Metadata zijn nodig voor bijvoorbeeld de identificatie (is dit het unieke object?), locatie (waar staat het object en wat zegt dit over de structuur?), beschrijving (wat is het, door wie is het waar, wanneer en waarom gemaakt?), leesbaarheid (welk formaat heeft het object?) en het rechtenbeheer (hoe zit het met de rechten, gebruiksvoorwaarden en restricties?).⁶⁶ Er zijn verschillende vormen van metadata, zoals bijvoorbeeld beschrijvende (bijvoorbeeld: auteur, titel, datum) en technische (bijvoorbeeld: bestandsnaam, wanneer het bestand is aangemaakt of gewijzigd, het formaat, de folder structuur, de gebruikte hardware/software) metadata.⁶⁷ Hoe meer je ervan verzamelt, hoe beter.

⁶⁶ Richtlijnen voor de selectie van digitaal erfgoed voor langetermijnconservering, pagina 18-19, door: UNESCO/PERSIST Content Task Force (2016). https://www.unesco.nl/sites/default/files/uploads/Comm_Info/finalpersistcontentguidelinesfinal_dutch.pdf

⁶⁷ Lees meer over metadata op <http://wiki.ncdd.nl/index.php?title=Duurzaamheidsbeleid/Beleidsuitgangspunt-Metadata>, op <http://www.den.nl/thema/37/metadata> en bij het National Institute Standards Organisation (NISO) (2004). Understanding Metadata. NISO press, USA (Online). <http://www.niso.org/publications/press/UnderstandingMetadata.pdf> (November 2011).

6.6 Stel je vondsten veilig

Om je vondsten veilig te stellen voor toekomstig gebruik heb je hulp van technische webarcheologen nodig. Het gaat hierbij niet alleen om het overzetten van data, maar ook: Hoe stel je je vondsten veilig met behoud van de metadata? Door simpelweg de bestanden te kopiëren loop je het gevaar dat je je metadata verliest.⁶⁸ In de volgende technische verdieping leggen we in verschillende stappen uit hoe je je data en metadata veilig kunt stellen.

Technische verdieping: Overzetten en veiligstellen van data

We beschrijven hier drie processen om alle data goed over te zetten en veilig te stellen voor toekomstig gebruik. 1. Zet je data over van de datadrager: Maak images. 2. Valideer de data: Maak checksums van de images en vergelijk ze. En tot slot: 3. Maak meerdere back-ups van alle data!

1. Zet je data over van de datadrager: Maak images

Let op! Voordat je enthousiast begint met het verzamelen van data, en de data klaarzet voor de overdracht naar een andere datadrager, moet je er zeker van zijn dat er zorgvuldig omgegaan wordt met de data. Het belangrijkste is dat er geen data verloren gaan. Maak daarom 'raw disk images' van de data. Let altijd op: Plaats nooit zomaar de datadrager in een machine, er bestaat een kans dat de data onbedoeld worden vernietigd.⁶⁹ Zorg dat, indien mogelijk, de data alleen gelezen kunnen worden (read-only). Als de datadrager read-only 'gemount' kan worden (al dan niet fysiek afgedwongen met een hardware switch), is dit natuurlijk het beste.

⁶⁸ Waarom kun je niet gewoon kopiëren en plakken? Voor de inhoud van de bestanden zelf werkt dit zonder problemen, maar er verandert iets aan de metadata. Wat precies hangt af van de operating system waarop je kopieert. Zo kan bijvoorbeeld de 'date modified' veranderen, maar ook de 'date created'. Ook gaat vaak informatie verloren waar je wellicht niet aan denkt, zoals bijvoorbeeld extended bestandsattributen. Deze kunnen cruciaal blijken, en zijn van historisch belang. Verder is informatie aanwezig in het bestandssysteem in de vorm van een swap-file, of gewiste bestanden op de computer. Als je de bestanden dus één op één overzet, kan al deze informatie verloren gaan.

⁶⁹ Je wilt voorkomen dat het systeem waarop jij de data wilt overzetten, automatisch het opslagmedium mount en runt. Zo kan alleen al het aanzetten van een systeem de data al veranderen. Er kan bijvoorbeeld een opstart proces zijn dat bestanden aanpast, of het OS wilt het bestandssysteem converteren naar een nieuwere versie. Of het OS draait automatisch een scandisk / fsck omdat het te lang geleden is dat het bestandssysteem werd gecontroleerd. Dit geldt ook voor virusscanners die mogelijk data weggooien. Je wilt dus een aparte, gecontroleerde machine gebruiken voor het overzetten, die nooit automatisch operaties doet op mediadragers. Deze machine gebruik je dus als een overdrager tussen de bron en het opslagmedium naar keuze. Ook is het mogelijk om een bootable Linux CD of USB te gebruiken om de data over te zetten vanaf de originele machine. Let wel op dat je meteen de eerste keer bij het opstarten dan vanaf de CD of USB opstart om te voorkomen dat het origineel verandert door bijvoorbeeld het opstart proces.

De beste manier om bestanden over te zetten is om een complete kopie te maken van het gehele opslagapparaat door deze te kopiëren als een ‘image’.⁷⁰ Je kunt op deze manier de bestanden die op een datadrager staan overzetten naar een ander opslagmedium. Het maken van een disc image voorkomt dat er (meta)data verloren gaan. In de image zet je dus de data over, een exacte een-op-een kopie van de gehele datadrager: bit voor bit, inclusief de partities, boot manager en lege ruimtes. Op deze manier voorkom je problemen met compatibiliteit van filesystemen. Zet je een werkend systeem over, dan kun je deze met compatibele hardware meteen weer aan de praat krijgen. Een image is dus een archiefbestand met alle data van het gekopieerde opslagmedium.⁷¹

2. Valideer de data: Maak checksums van de images en vergelijk ze

Om honderd procent zeker zijn dat de data die je van de originele datadrager afhaalt correct zijn, maak je meerdere images. Daarna maak je van alle images een checksum. Hiermee controleer je of het allemaal goed is gegaan. Een checksum genereert een reeks karakters, die altijd gelijk is voor dezelfde data. Als het maken van de images goed gegaan is levert dit dus iedere keer dezelfde reeks karakters op. Door de reeks karakters dus te vergelijken kun je controleren of de images hetzelfde zijn.

Hoe doe je dit? Maak met een tool zoals md5sum en sha1sum een cryptografische hash van de raw disk-image. Vergelijk deze hashes (de MD5 hashes, en de SHA hashes). Zijn alle MD5 hashes hetzelfde? Zo nee, dan was er een data-error bij het kopiëren van minimaal één van de pogingen. Let op: Bewaar de hashes als onderdeel van je documentatie. Ze zijn het bewijs voor een e-depot en toekomstige onderzoekers dat het oorspronkelijke bronmateriaal niet gewijzigd is. Als je de image maakt van het gehele systeem is het belangrijk dat je op dat moment ook de checksums maakt van het systeem. Dit kan bij een image niet achteraf gedaan worden, want de checksum verandert op het moment dat je de image weer benadert. Dit is ook een van de redenen dat je altijd moet werken met een kopie van de image, niet het origineel.

⁷⁰ Je kunt een image genereren via Unix command-line tools zoals ‘dd’ (of dd_rescue), of via forensics tools. Bij het gebruik van forensic tools kun je de data het beste opslaan in het Advanced Forensics Format (AFF4) aangezien dit formaat ook door andere tools breed ondersteund wordt. Houd er alleen wel rekening mee dat dit formaat niet zomaar gebruikt kan worden in een virtual machine, of zomaar gemount kan worden met een ‘mount -o loop / disk.img /path/’ onder linux. Bij het gebruik van ‘dd’ onder linux is het verder nog van belang te weten dat het noodzakelijk is om de optie: ‘conv=noerror’ mee te geven. Waarom is dit van belang? Als er block errors zijn, dan dienen deze opgevuld te worden met nullen, aangezien anders de posities van blocks veranderd, en het de totale boekhouding van het filesystem niet meer klopt.

⁷¹ Een alternatief voor een image is het maken van een archiefbestand van de data. Dit doe je bijvoorbeeld als het niet met een complex systeem gaat, maar om alleen een paar bestanden handelt. Het veelgebruikte tar format kan dit bijvoorbeeld. Maak dan alleen een checksum van het bestand, en niet van het hele opslagmedium. Je behoudt dan wel de interne metadata, maar je verliest wel een de context van het systeem waarop de data draaiden.

3. Maak meerdere back-ups van alle data!

Wanneer data belangrijk zijn willen we zeker weten dat ze ook beschikbaar blijven. 'It is better to be safe than sorry'! Maak daarom met regelmaat kopieën van de belangrijke data. We raden de '3-2-1 back-up' strategie⁷² aan. Dit is een simpele set van drie regels waaraan voldaan moet worden om een goede back-up te hebben.

- Zorg voor minimaal drie kopieën van de data. Hoe meer kopieën hoe beter natuurlijk.
- Gebruik minimaal twee verschillende methoden van opslag. Hiermee bedoelen we verschillende opslagmedia. Door verschillende opslagmedia te gebruiken, zoals bijvoorbeeld een externe harde schijf, SSD, USB-stick, CD, DVD, DLT-tape, SD-kaart of een cloud storage, verklein je de kans dat een type opslag niet meer werkt. Vergeet in het laatste geval niet encryptie te gebruiken bij privacygevoelige informatie.
- Zorg dat minstens een van de back-ups op een andere fysieke locatie staat. Zo voorkom je dataverlies als er bijvoorbeeld brand uitbreekt op een locatie.
- Bij het maken van een back-up geldt hetzelfde als bij het maken van een image: Controleer al je stappen door het maken van checksums.

Technische verdieping: Verspreid je risico's

Je kunt nog meer veiligheid nastreven door 'redundancy' (meervoudigheid) toe te voegen. Door een RAID ('Redundant Array of Independent Disks') opstelling in te zetten verspreid je risico's. Hierbij gebruik je meerdere harde schijven om de data beschikbaar te houden, ook bij uitval van een schijf. Let wel: RAID is niet hetzelfde als een back-up. Hoewel het beschermt tegen uitval van een schijf, beschermt het niet tegen user error of corruptie van het filesystem. Je krijgt bij (langetermijn) opslag altijd te maken met bitrot. Daarom adviseren wij om gebruik te maken van een filesystem dat interne checksums heeft, zoals ZFS. Mensen gebruiken wel eens versioning systemen zoals bijvoorbeeld GIT of Apple Time Machine als back-up. Let op: Doe dit niet, want versioning systemen gebruiken altijd deduplicatie. Dit betekent dat corruptie in de data alle versies van de data zal aantasten.

⁷² Bekijk deze video over 'Data Storage 3-2-1': <https://www.youtube.com/watch?v=F5E7jDcZ30c>

07

Reconstrueren

Hoe reconstrueer je digitale objecten?

In deze fase ga je aan de slag met de gevonden en veiliggestelde data. Je mag nu de schatkamer in en de pakketjes uitpakken. In dit hoofdstuk kijk je naar de verschillende processen, technieken en methoden om born digital materiaal te onderzoeken en het digitale object te herstellen en/of na te bouwen. Door de digitale objecten weer werkend te maken draag je bij aan het behoud van het digitale erfgoed. En op deze manier kun je inzicht verwerven in de werking, het gebruik en betekenis ervan, en kun je als het lukt het publiek de geschiedenis laten (her)beleven. Verschillende duurzaamheidsstrategieën (migratie, conversie, emulatie of herinterpretatie van de data) passeerden al de revue. Een aantal hiervan wordt in dit hoofdstuk verder uitgediept. Mogelijk is jouw project niet gericht op reconstructie, maar alleen op permanente archivering. In dit geval kun je dit hoofdstuk overslaan. Voor iedereen geldt: Wees je ervan bewust dat de keuzes die je maakt de mogelijkheden van de toekomstige onderzoekers kunnen bepalen (inperken). Ga op ontdekkingstocht en start je reconstructiewerkzaamheden.

Stappenplan:

- 1 Beschrijf de uitgangspunten
- 2 Onderzoek de ethische kwesties
- 3 Start met de voorbereidende werkzaamheden
- 4 Reconstrueer: emuleren of restaureren?
- 5 Duid de data

7.1 Beschrijf de uitgangspunten

Je hebt je doelstellingen geformuleerd en je hebt inmiddels je objecten gevonden. De volgende stap is: Wat wil je dat ermee gebeurt? Om richting te geven aan de zoektocht in de opgegraven en veilig gestelde data, en om een kader te bieden aan de hand waarvan het team kan gaan reconstrueren, stel je uitgangspunten op.⁷³

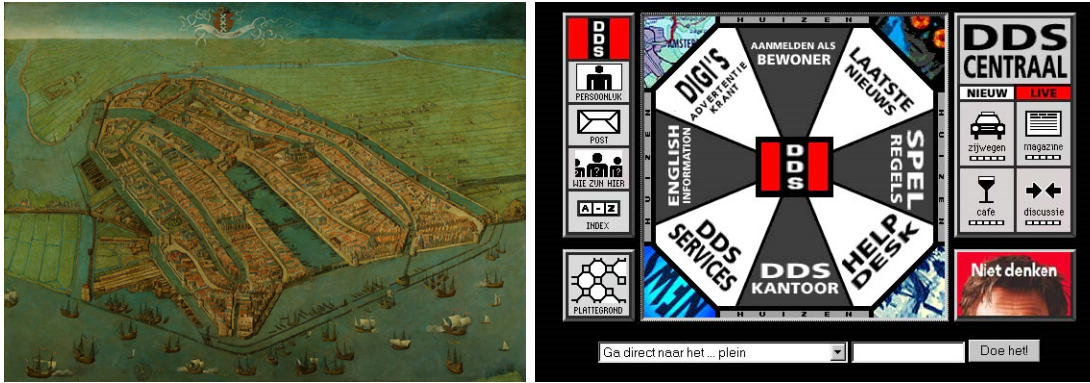
Onze aanpak: case study 'DDS Herleeft'

DDS is een aansprekend voorbeeld van vroege internetgeschiedenis. Door de verschillende varianten van de virtuele stad werkend te laten zien maak je het mogelijk dat mensen de ontwikkeling van het internet en de enorme sprongen die deze ontwikkeling heeft doorgemaakt kunnen ervaren.

Onze uitgangspunten:

- Het restaureren, conserveren, behouden en presenteren van digitale objecten zijn nieuwe en belangrijke taken van musea. Het reconstrueren en duurzaam ontsluiten van (born) digital objecten stelt andere eisen dan fysieke objecten dat doen. Hoe pak je dit aan? Hoe zorg je ervoor dat bedreigde of verloren gegane digitale objecten (weer) beschikbaar worden en blijven voor musea?
- Door het digitale object weer werkend te maken kan het publiek weer door De Digitale Stad wandelen en de evolutie van deze unieke virtuele stad en het vroege web ervaren en (her)beleven. Hoe kun je de oorspronkelijke gebruikerservaring weer terugbrengen? Hoe repareer je dynamische webobjecten die interactief, vernetwerkt, procesgericht en contextafhankelijk zijn? Hoe reconstrueer je historische data in het spanningsveld tussen 'historisch getrouw' (authentiek) zijn en doen wat haalbaar is binnen de mogelijkheden ('quick & dirty')?
- Als webarcheoloog ben je verantwoordelijk voor de juridische en ethische aspecten. Hoe ga je onder andere om met de privacy van de DDS-bewoners van toen (let wel, het web is nog jong, de meeste betrokkenen leven nog!).

⁷³ Laat je inspireren door het boek 'Preserving Complex Digital Objects' door Janet Delve, David Anderson (editors). Facet Publishing, 2014.



Twee unieke stadsgezichten, een fysiek en een born digital object: 'Gezicht op Amsterdam in vogelvlucht' (1538) door Cornelis Anthonisz. en DDS3.0 (1994)

7.2 Onderzoek de ethische kwesties

Zoals al eerder aangestipt vormt het onderzoek naar de juridische en ethische aspecten een belangrijk onderdeel in het omgaan met digitaal erfgoed. Sterker nog, het is de verantwoordelijkheid van de webarcheoloog. Je moet aantonen dat je met gepaste zorgvuldigheid handelt en je uiterste best doet en alle inspanningen levert, waar mogelijk, om de rechten van alle betrokkenen te waarborgen. We gaan in deze paragraaf, voordat we aan de slag gaan met het reconstrueren, dieper in op de ethische kwestie van privacy.

TIP: Markeer de rode lijnen

Door in een vroeg stadium ethische spanningsvelden te verkennen kun je de rode lijnen waar je niet overheen wilt gaan al vastleggen voordat je daadwerkelijk met de data aan de slag gaat. Dit kan vervolgens zijn weerslag krijgen in verwerkers-contracten, maar het is vooral van belang dat het totale team waarmee je samenwerkt op de hoogte is van deze rode lijnen. Deze rode lijnen zijn idealiter de uitkomst van een brainstorm van het totale team. Het staat natuurlijk buiten kijf dat deze 'ethische box' zich binnen de dwingende box van de wetgeving bevindt (check dat dus eerst).

Onze aanpak: case study 'DDS Herleeft'

Voordat we aan de slag zijn gegaan met het verkennen van de data en de reconstructie van de virtuele stad, stelden we onszelf een aantal vragen. Hoe zit het met de grenzen tussen de publiek toegankelijke informatie (bijvoorbeeld de algemene informatie over DDS en het nieuws) en de privégegevens van de gebruikers (bijvoorbeeld de persoonsgegevens en de e-mails)? Welke rol en verantwoordelijkheden hebben de nieuwe beheerders (de huidige webarcheologen)? Hoe kunnen we recht doen aan de bouwers en bewoners van toen (en nu! de meeste betrokkenen leven nog!)? Met deze vragen zijn we op onderzoek gegaan.

Enkele weken na de 'Grave Diggers Party' kregen wij, zoals we eerder beschreven, een hele bijzondere vondst in handen: De FREEZE. Uit gesprekken met de oud-bewoners en door onderzoek in het Internet Archive wisten we inmiddels echter ook dat er stevig protest werd aangetekend tegen het maken van De FREEZE.

De bouwers van DDS wilden de bijzondere geschiedenis van De Digitale Stad bewaren, want zij zagen het als volgt: "De FREEZE is niet meer dan een symbolische poging om aan de ene kant te laten zien dat dit medium erg vluchtig is, dat er constant naar de toekomst wordt gekeken, terwijl er ondertussen ook geschiedenis wordt gemaakt".⁷⁴ En het opslaan van e-mail was "een unieke mogelijkheid voor toekomstige historici op een indruk te krijgen van taal, gebruiken, ideeën en leven van de Eerste Digitale Stad Bewoners".⁷⁵ Bij bewoners riep het maken van De FREEZE juist vragen op over hun privacy. In de discussiegroep dds.dds, op De FREEZE pagina in de HyperNews discussie vooraf aan het maken van De FREEZE en tijdens de bewonersbijeenkomst op 13 januari 1996, werd er heftig gediscussieerd. Een aantal bewoners maakte duidelijk dat het bewaren en op den duur beperkt toegankelijk maken van hun privégegevens niet op prijs werd gesteld. Aanvankelijk wilde DDS een volledige freeze maken van alle data op de centrale computers, inclusief de persoonsgegevens en persoonlijke e-mail. Maar omdat DDS medewerkers de bewoners wilden tegemoet komen en hun privacy niet in gevaar wilden brengen, werd er uiteindelijk voor gekozen om alleen de publieke informatie te bewaren. Alleen bestanden die publiekelijk opvraagbaar waren, zoals homepages, gebouwen, discussiegroepen, cafés, IRC activiteiten etc., zouden in De FREEZE bewaard blijven. De privégegevens van de gebruikers (persoonsgegevens, e-mail en persoonlijk bestanden) zouden niet bewaard worden.

⁷⁴ Bron: <http://web.archive.org/web/19991110021440/http://www.dds.nl:80/HyperNews/get.cgi/freeze/8/1/1/1/1.html>

⁷⁵ Bron: <http://web.archive.org/web/19991109020934/http://www.dds.nl:80/dds/nieuws/1-96.html>

Uiteindelijk werd gekozen voor de volgende strategie: Een week na De FREEZE werd een webformulier online gezet (van 24 januari tot 15 februari 1996), waarin bewoners zich vrijwillig konden aanmelden zodat hun persoonlijke gegevens wel bewaard zouden blijven voor de toekomstige historici. Er zouden voorwaarden worden geformuleerd voor het toegang krijgen tot die privégegevens en deze zouden met enkele deskundigen worden bekeken.⁷⁶ In het formulier konden bewoners die dat wilden toestemming geven. "Ja, ik geef toestemming om mijn persoonlijke e-mail en gegevens in De FREEZE op te nemen." In het 'Freeze Privacy-formulier' stond: "De FREEZE is de backup die op 15 januari 1996, de tweede verjaardag van De Digitale Stad, gemaakt is. De gegevens mogen pas na 50 jaar openbaar gemaakt worden."⁷⁷

Bij het reconstrueren van DDS besloten wij, waar mogelijk, zoveel mogelijk rekening te houden met zoals we het in de huidige debatten kennen het recht om vergeten en herinnerd te worden. Onderdeel van ons komend onderzoek was het vinden van de aanmeldingen van het 'Freeze Privacy-formulier'. Een andere stap was het opstellen van een geheimhoudingsovereenkomst voor alle betrokken webarcheologen.⁷⁸ Zonder dit contract te tekenen kregen betrokkenen geen toegang tot de data. Bij ondertekening verklaarden ze onder andere niets zonder toestemming te publiceren en na onderzoek de (eventueel opgeslagen) data te verwijderen.

⁷⁶ Bron: <http://web.archive.org/web/19991109020934/http://www.dds.nl:80/dds/nieuws/1-96.html>

⁷⁷ Bron: <https://web.archive.org/web/19970618011433/www.dds.nl/dds/nieuws/freeze/form.html>

⁷⁸ Voorbeelden van een geheimhoudingsovereenkomst en een bewerkersovereenkomst zijn te vinden op <https://www.nba.nl/tools-en-voorbeelden/model-bewerkersovereenkomst/> en <https://blog.legaldutch.nl/geheimhoudingsverklaring-nda-opstellen/>

7.3 Start met de voorbereidende werkzaamheden

Tijd voor de voorbereidende werkzaamheden om de reconstructie mogelijk te maken. Wat heb je hiervoor nodig?

Stappenplan:

- Stel je team samen
- Maak data toegankelijk
- Verken de data

Stel je team samen

Om aan de slag te kunnen gaan heb je experts nodig die én kennis hebben over het object dat je wilt reconstrueren, én de gereedschapskist met instrumenten kunnen bedienen waarmee je het object kunt reconstrueren. Als je de kennis en vaardigheden zelf niet in huis hebt is het tijd om een partner in te schakelen. In het plannen van je project heb je de kennisgebieden al in kaart gebracht en heb je waarschijnlijk al in je netwerk rondgekeken en potentiële partners gevonden. Naast de nieuwe webarcheologen met wie je gaat samenwerken wil je ook, als dit mogelijk is, de oorspronkelijke ontwikkelaars van je object betrekken. Zij zijn een bron van kennis en ervaring en kunnen uitleg geven over de historische context van het object. Zij kunnen de huidige experts leren om niet alleen uit te gaan van de hedendaagse praktijken zoals zij die kennen en toepassen, maar ook leren terugdenken in (en reflecteren op) de tijd dat er op een andere manier werd ontwikkeld.

Onze aanpak: case study ‘DDS Herleeft’

Wij hadden niet alle kennis of onbeperkte mankracht in huis. Onze technische experts werkten op vrijwillige basis en konden slechts beperkte tijd vrijmaken. Na een periode van weinig vooruitgang besloten wij dat het tijd was om naar de Universiteit van Amsterdam te gaan. Met onze wensen, inspiratie en harde schijven vol opgegraven data meldden we ons. We hadden ook in kaart gebracht waar de nieuwe technische webarcheologen rekening mee moesten houden. Ten eerste was dat de kennis van het object (DDS) dat ze zouden gaan bestuderen en reconstrueren: Perl, C, Bash (shell script), Javascript, HTML, SUN OS, SPARC. En ten tweede bestond dit uit kennis van de instrumenten waarmee ze zouden onderzoeken en reconstrueren: UNIX, tools voor visualisatie, geolocatie, natural language processing en doorzoekbaar maken en filteren.

Universitair hoofddocent Gerard Alberts stelde zijn college ‘History of Digital Cultures’⁷⁹ open, en een flink aantal slimme studenten informatica van de Master of Computer Sciences pakte de uitdaging op. Eerst moesten de docent, docent-assistenten en studenten een contract (geheimhoudingsverklaring) tekenen. Daarna konden ze ‘met de poten in de modder’ (de kleilaag uit het vroege web) en aan de slag met de opgegraven (rauwe) data.

Van 2015 tot en met 2017 zijn meerdere studenten in de verschillende lagen klei (data) afgedaald. Eerst verkennen, documentatie lezen, brokstukken onderzoeken, de context bestuderen, en daarna voorzichtig aan de hand van de overblijfselen uit het verleden probeerden zij DDS te reconstrueren en te interpreteren. Doel was: Inzicht bieden in de (bestaande en nieuwe) processen, technieken en methoden om born digital materiaal en de context waarin zij worden aangetroffen, op te graven en te reconstrueren. Naast de techniek hebben de studenten ook de historische en juridische aspecten onderzocht.

Gerard Alberts organiseerde verschillende bijeenkomsten met oud-medewerkers van DDS. Deze ontmoetingen tussen de verschillende generaties programmeurs en systeembeheerders waren bijzondere aangelegenheden en gaven de studenten de mogelijkheid om vragen te stellen: hoe werkten de DDS en alle onderdelen, wat was de bedoeling en waarom hebben jullie het zo gedaan?

De studenten legden hun werkzaamheden vast in rapporten en hebben een aantal keer hun werk gepresenteerd. In een wetenschappelijke context bij de Universiteit van Amsterdam, en in het Amsterdam Museum en bij Waag Society waarbij geïnteresseerden aanwezig waren, waaronder oud-medewerkers en oud-bewoners.⁸⁰ Voor de studenten waren dit unieke gelegenheden om oog in oog te komen met het 'object' van hun studie: Wie waren de oorspronkelijke bewoners, wat voor ideeën hadden ze en vooral: wat deden ze?

Maak data toegankelijk

Om met de data aan de slag te gaan moet je er natuurlijk wel bij kunnen. Bij een kleinschalig project kun je in de beginfase de data op een (beveiligde) computer zetten en met de data werken. Wanneer je te maken hebt met grotere teams of verschillende organisaties kun je niet om een schaalbare oplossing heen. Denk hierbij aan een goede (besloten) werkomgeving. En, werk altijd alleen met een kopie van de data, bijvoorbeeld van één van de images die je veilig hebt gesteld, zoals beschreven in het vorige hoofdstuk.

⁷⁹ Het college 'History of Digital Cultures' van dr. Gerard Alberts maakt onderdeel uit van de 'Master's in Computer Science' aan de Universiteit van Amsterdam. In dit college verwerven studenten historische kennis over 'computing' in haar culturele context. Ze verdiepen zich in de theorieën over de informatiemaatschappij en digitale cultuur. Door de ervaring met een ander vakgebied (geschiedenis) leren de informatica studenten te reflecteren op hun eigen discipline. Naast theorie en contemplatie leren de studenten werken in teamverband. Zie: <http://studiegids.uva.nl/xmlpages/page/2016-2017/zoek-vak/vak/28352>

⁸⁰ De resultaten van de studenten zullen worden gedeeld op het communicatieplatform van het museum: <https://hart.amsterdam/re-dds>.

Onze aanpak: case study 'DDS Herleeft'

In het eerste jaar is een klein team studenten aan de slag gegaan met de data. Hiervoor gebruikten ze een kopie van De FREEZE op een USB-stick. Vanaf de USB-stick zetten zij een kopie op hun eigen machines en konden zo werken. Deze simpele en informele aanpak werkte voor de kleine groep. Voor het tweede jaar van het college is een server opgezet zodat de data, resultaten en documentatie gedeeld konden worden met het hele team. De studenten maakten op afstand een verbinding met de server (achter een login!) en deelden zo hun data en hun werk.

Technische verdieping: Zet een werkomgeving op

Om effectief te kunnen werken tijdens het college History of Digital Cultures is een ontwikkelserver opgezet met de benodigde data. Dit was een Linux-server in de vorm van een Intel NUC. Iedere student kreeg een eigen account, dat na het tekenen van de geheimhoudingsovereenkomst werd geactiveerd op de server. Dit waren geen admin accounts. De studenten konden bijvoorbeeld geen programma's installeren. Het inperken van de rechten was puur om te voorkomen dat (slimme) studenten, wel of niet kwaadwillend, de server plat konden leggen. Studenten logden op de server in met een encrypted ssh verbinding. Op de server was een read-only map met de te onderzoeken data. Dit was om ruimte te besparen. In plaats van iedere student een eigen kopie te geven, deelden ze een kopie die ze niet konden aanpassen. De studenten konden een bestand kopiëren naar hun eigen home folder of eigen machine, en werkten daar met de data.

Afhankelijk van het materiaal, het doel en vaardigheden van het onderzoeksteam werden tools beschikbaar gesteld. Voor het college hadden de studenten toegang tot een SQL-database op de server. Voor elk team was een eigen database aangemaakt. De teams hadden leesrechten op alle databases en alleen edit-rechten op hun eigen database. Hierdoor konden ze gebruik maken van het werk van andere teams, maar konden ze het niet veranderen. Ook was er een Apache-webserver opgezet waar de studenten webpagina's konden publiceren om het onderzoek makkelijk te kunnen delen met de andere teams. Deze webpagina's waren afgeschermd van de buitenwereld door middel van http basic auth; want de mogelijkheid bestond dat er privégegevens bij zaten. Wanneer er extra tools nodig waren voor de studenten moesten zij dit eerst aanvragen bij de systeembeheerder. Aan het einde van het college werden de accounts van de studenten op non-actief gezet waardoor zij geen toegang meer hadden tot de server en verwijderden ze de data van hun computers, zoals afgesproken.

Verken de data

De data zijn binnen, bereikbaar en klaar om ontsloten te worden! ...en nu? Tijd om een paar stappen terug te doen en te bekijken wat je nu eigenlijk hebt. Blindelings de data in duiken is net als een bibliotheek in stappen zonder te weten waar de boeken staan of hoe ze georganiseerd zijn. Duik samen met onze technici in de DDS-data in de komende paragraaf.

Onze aanpak: case study 'DDS Herleeft'

Technische verdieping: Doorzoek de digitale hooiberg

Aangezien er in de afgelopen twintig jaar veel veranderd is, was het erg nuttig dat de kennis van de oud-DDS developers en systeembeheerders aangesproken kon worden. Een voordeel van een jonge discipline als webarcheologie is dat de bedenkers en bouwers veelal nog rondlopen. Maak hier gebruik van.

Zoals een student beschreef: "Door de 'oral history' en alle informatie die inmiddels verzameld was, wisten we waarmee we werkten. We hadden een kopie van de drie servers waar DDS op draaide: DDS, Alibaba en Shaman. We wisten dat we met een webserver te maken hadden. We weten hoe die er tegenwoordig uitzien, dus we konden zoeken naar soortgelijke structuren die destijds ook gebruikt werden. Zeker tijdens het pionieren in de vroege web-jaren werden er veel creatieve oplossingen gebruikt die anders zijn dan de huidige normen. Samen met de toenmalige systeembeheerders konden we de missende stukken invullen, waardoor we een goed beeld kregen van de structuur van de data."

'Find' en 'grep'

Als je weet waar de data ongeveer staan, kun je gericht gaan zoeken. Maar uitzoeken in welke hooiberg de data zich bevinden is bewerkelijk. We halen daarom de metaforische stoffertjes van de webarcheoloog uit de kast. Voor het zoeken van bestanden zijn er twee Unix command-line tools die van grote waarde kunnen zijn, namelijk 'find' en 'grep'. Find⁸¹ helpt je zoeken naar bestanden met grote precisie. Zo kun je zoeken naar bestandsnamen, maar ook reguliere expressies gebruiken. Je kunt verder bestanden vinden die tussen specifieke datums aangemaakt of gewijzigd zijn. Je kunt ook bestanden vinden die tot een bepaalde user behoren. Met 'grep'⁸² zoek je *in* bestanden. Zo kun je de inhoud van bestanden automatisch doorzoeken op zoektermen. Je kunt de tools ook combineren, dat levert nog meer precisie op.

⁸¹ Find (GNU findutils) 4.4.2 - Linux General Commands Manual, Eric B. Decker, James Youngman, and Kevin Dalley.

⁸² Grep (GNU grep) 2.20 - Linux General Commands Manual, Mike Haertel et al.

Metadata dating

We hebben al eerder geschreven over het belang van metadata. Aan de hand van de metadata weet je wanneer een bestand is aangemaakt en wanneer deze voor het laatst gewijzigd is. Met deze stukjes metadata kun je dus iets zeggen over de data zelf, wij noemen dit 'metadata dating'. Vernoemd naar 'carbon dating', een methode om de leeftijd van fossielen te bepalen. Zo kun je ook de leeftijd van jouw databron bepalen. Metadata dating functioneert ook als een belangrijke tool om te controleren of de data zijn wat je verwacht. Zo zou je in De FREEZE geen bestanden verwachten die gemaakt of gewijzigd zijn na 15 januari 1996. Als dat wel het geval is, dan is er iets fout gegaan met de metadata van De FREEZE. Met metadata dating kun je belangrijke informatie over de ontwikkeling van het digitale object terugvinden. Zo kun je bijvoorbeeld kijken naar wanneer e-mailbestanden werden aangemaakt. Was dit doordeweeks of juist in het weekend? Werden er veel bestanden aangemaakt in een bepaalde periode of juist niet?

DeNISTing

Nu ga je het kaf van het koren scheiden en de data analyseren. Eén van de standaard tools die je meteen kunt gebruiken is DeNISTing. We laten in deze paragraaf zien hoe wij DeNISTing hebben gebruikt én hebben aangevuld. Deze aanvulling noemen wij DeNISTing-NG: de Next Generation.

Als onderdeel van onze gereedschapskist om de bestanden te herkennen en te duiden zijn we met DeNISTing aan de slag gegaan. Dit is een van de methoden die we geleend hebben van de digitale forensics en e-discovery wereld.⁸³ Om uit te leggen wat dit is, moeten we even de techniek in duiken. Er bestaat iets wat men een 'cryptografische hash' noemt. Een cryptografische hash van willekeurige data haalt deze data zodanig door elkaar dat alle bits in de data 'gecondenseerd' worden tot één enkel getal van een bepaalde lengte. Alle bits dragen hierin bij aan de eindwaarde van dit getal, dan typisch 256 bits lang is. Als een stukje data dus ook maar één enkel beetje anders is dan een ander stukje data, dan zal de hash van deze data er totaal anders uit zien. Als je dit doet bij bestanden, dan krijgt elk bestand dus een uniek nummer. Een hash van een bestand is een 'digitale vingerafdruk'.

Het gebruik van hash'es maakt het dus mogelijk om bestanden te herkennen aan hun eigen unieke vingerafdruk. Als je als advocaat in de Verenigde Staten of als rechercheur een computer wilt analyseren, dan ben je niet geïnteresseerd in al die bestanden die van het besturingssysteem of van standaard software komen. Je wilt alleen

⁸³ Bekijk de gratis tool voor 'digital forensics' en digitale preserving: Bitcurator. <http://www.bitcurator.net/>.

die bestanden bekijken die niet bij standaard software hoort. Maar hoe scheid je nu al die niet-relevante bestanden van de relevante data? Zou het niet fijn zijn als er een ongelooflijk grote database was van bijna alle software die ooit gemaakt was, met daarbij voor elk bestand de unieke vingerafdruk? Die bestaat! Het NIST ('National Institute of Standards and Technology') in de Verenigde Staten is al geruime tijd bezig met het archiveren van software vanaf het moment dat deze gepubliceerd wordt. Het doel van het verzamelen van alle software is dat je het kunt inzetten om oude bestanden in de toekomst uit te kunnen lezen. Dit maakte het mogelijk voor het NIST om ook een database aan te leggen met alle vingerafdrukken van alle bestanden in deze software. Dit wordt in de e-discovery / forensics wereld de 'NIST hashset' genoemd. Deze hashset is gratis te downloaden, en kan gebruikt worden om alle niet-relevante software bestanden weg te filteren. Dit proces heet DeNISTing. Het is als het vinden van de speld in de hooiberg door het wegscheppen van al het hooi, zodat alleen de speld overblijft.

DeNISTing-NG

Wij zijn verder gegaan met DeNISTing als concept. We hebben bedacht dat je niet alleen de NIST hashset kunt gebruiken om software bestanden weg te scheppen, maar ook om software te herkennen. Op het moment dat je bijvoorbeeld de hashset loslaat op een harde schijf, kun je zien dat: 1. Het systeem Windows 3.11 gebruikt als operating system, 2. WordPerfect 5.1 op het systeem staat en 3. Lotus Notes gebruikt wordt als e-mail client. Je kunt dus een lijstje krijgen van alle software op het systeem.

Bovendien kun je meer met deze zelf ontwikkelde software. Je kunt ook zien dat het systeem nieuwer moet zijn dan jaar X, aangezien de software pas in jaar X uitgegeven werd. Je kunt bestanden die niet herkend werden door DeNISTing, maar wel aanwezig zijn in de directory van software die wel herkend is, makkelijker duiden. Een bestand in een WordPerfect directory met een doc extensie is waarschijnlijk een WordPerfect bestand. Helemaal leuk wordt het als je meerdere back-ups van hetzelfde systeem hebt. Je kunt dan dankzij de software upgrades die plaatsgevonden hebben, zien wat waarschijnlijk de oudste, nieuwere en nieuwste versie is van de back-up van het systeem in kwestie. Zonder metadata kun je dus toch -volledig geautomatiseerd- iets zeggen over een collectie bestanden. Dit noemen wij DeNISTing-NG (Next Generation).

We zijn ook bezig om gevonden bestands-collecties automatisch in de juiste volgorde te plaatsen, en waar mogelijk samen te voegen. Dit betekent dat als een bestand in back-up één aanwezig is, en ook in back-up drie, dat het waarschijnlijk is dat deze ook in back-up twee

thuishoort. Als back-up twee incompleet is, dan kun je deduceren dat deze waarschijnlijk ook in back-up twee aanwezig zou zijn als deze ook compleet zou zijn. Kortom, je kunt hiermee verschillende bestandssystemen van hetzelfde systeem in de juiste volgende plaatsen, en missende bestanden aanvullen. Vermeldenswaardig is het gebruik van zogenoemde 'Fuzzy Hashes'. Dit zijn hashes die, in tegenstelling tot normale cryptografische hashes, proberen om juist wél zoveel mogelijk informatie van een bestand te behouden (via 'rolling hashes'). Deze techniek wordt door analisten die virussen classificeren gebruikt om virussen die een klein beetje anders zijn toch te herkennen. Dit is ook in te zetten om databestanden die vergelijkbaar zijn te herkennen. Deze tools bij elkaar maken het mogelijk om volledig geautomatiseerd complete sets van bestanden te herkennen en in context en volgorde (qua tijd) te plaatsen. Door deze met een valse datum in een software versioning control systeem te zetten, is het vervolgens mogelijk om een 'point-in-time' check-out van alle bestanden te doen, die op dat moment op het filesysteem aanwezig zouden moeten zijn. Dit laatste systeem noemen wij een 4D-collectie, aangezien tijd ook als factor aanwezig is.

We hebben helaas geen 4D-collectie kunnen maken, omdat we maar één werkbare dataset hadden, te weten De FREEZE. Verder is de NIST hashset sterk gericht op de meest gebruikte besturingssystemen, zoals Windows, terwijl wij te maken hadden met SunOS en zelf gecompileerde GNU tools.

TIP: Gebruik de DeNISTing-NG tools

De ontwikkelde DeNISTing-NG tools zijn te vinden op: <https://github.com/tifkap/>. Dit is een 'work in progress', dus het gebruik kan lastig zijn. Contacteer Paul Vogel voor details.⁸⁴

⁸⁴ Paul Vogel: spam@wateenlangedomeinnaamdieisvastnietecht.nl (systeembeheerder, programmeur).

7.4 Reconstrueer: emuleren of restaureren?

Als alle data ontsloten zijn en de uitgangspunten zijn vastgesteld onderzoek je de mogelijkheden. Een oorspronkelijk systeem weer werkend te krijgen heeft grote historische waarde en betekenis. Hoe 'historisch getrouw' (authentiek) je het digitale object kunt herstellen is afhankelijk van je teruggevonden data, hard- en software. Is een emulatie mogelijk? Is een restauratie noodzakelijk? Is een replicatie wenselijk? Een emulatie is bit voor bit en byte voor byte gelijk aan het origineel. Een restauratie maakt zoveel mogelijk gebruik van de originele programma's en data en past de componenten (of data) aan om het digitale object weer werkend te krijgen. Een replicatie is een geheel nieuw opgebouwd object met dezelfde functionaliteit als het origineel, al dan niet met de originele data.

Emuleren is altijd meer authentiek dan een restauratie. Een replicatie heeft vaak alleen authentieke data. Vergelijk het met een Amsterdams grachtenpand. Als je de keuze hebt tussen het bewaren van het oorspronkelijke pand op een nieuwe fundering (emulatie), of het herbouwen van het pand met behoud van alleen de oude gevel (restauratie), dan heeft het behoud van het gehele oude pand meer historische waarde. Je kunt ook een pand weer helemaal opnieuw bouwen met nieuwe materialen, maar met behoud van de originele meubelen (replicatie). Het eindresultaat mag er dan van buitenaf hetzelfde uitzien, maar van binnen is het een wereld van verschil.

Technische verdieping: De vormen van reconstructie

In deze verdieping gaan we verder in op de voor- en nadelen van emuleren, restaureren en repliceren.

Emulatie

Bij emulatie wordt het oorspronkelijke (binaire) materiaal gebruikt. Dat betekent dat al het materiaal (machine code) ongewijzigd is. Als bijvoorbeeld de originele computer architectuur SPARC was, dan moet de reconstructie op een SPARC-emulator draaien. Een emulator is een systeem dat een ander systeem nadoet. In het geval van het grachtenpand is het originele pand intact en staat deze op een nieuw fundament (dit is de emulator).

Wanneer je IT'ers vraagt of iets mogelijk is met de computer, zeggen ze doorgaans zonder te twifelen: Ja! Het lijkt dan ook makkelijk om een opgegraven systeem onder een emulator aan de praat te krijgen. Hoe moeilijk kan het zijn om de volledige systemen onder een moderne emulator weer aan de praat te krijgen, in plaats van alles stukje voor stukje werkend te krijgen?

In principe zijn alle systemen in een emulator te draaien, maar in de praktijk is dit minder makkelijk dan het lijkt. Systemen staan namelijk zelden op zichzelf, en hun externe afhankelijkheden dienen dan ook geëmuleerd te worden. Ook zijn bouwstenen zoals oude versies van software vaak niet meer aanwezig, wat het lastig maakt om het hele

OS te emuleren. Daarnaast is het gebruik van DRM / licentie servers (al dan niet op het internet) een serieus obstakel. Deze services bestaan vaak niet meer. Vergelijkbare problemen bestaan bij OSX, die eist dat bepaalde, door Apple gecontroleerde chips aanwezig zijn om te voorkomen dat je OSX kunt draaien zonder hardware van Apple te kopen.

Een voordeel van emulatie is dat, als het mogelijk is, het meestal snel en makkelijk op te zetten is. Een nadeel is dat als je een andere architectuur moet emuleren dan de architectuur waar de emulator op draait, dit grote vertraging kan veroorzaken.

Restauratie

Bij restauratie worden de oude software en data opnieuw werkend gemaakt. Dit gaat alleen om de applicaties en de benodigdheden, en niet om het originele OS. Restauratie is een beperkte vorm van reconstructie. Voor zover mogelijk worden alle componenten van het originele systeem hergebruikt. Soms worden deze aangevuld met nieuwe componenten. En in het algemeen worden deze op een nieuw, modern systeem neergezet. Dit systeem kan een andere architectuur (en libraries) hebben. In dit geval, en indien de source code aanwezig is, dient de applicatie in het algemeen opnieuw gecompileerd te worden. Soms moeten hele stukken code aangepast worden om het compileren mogelijk te maken. Dit kan een moeilijk en tijdrovend proces zijn. Het komt vaak voor dat je een oude versie van een library moet vinden om überhaupt de applicatie te kunnen compileren. Naast de hierboven beschreven problemen om de server applicaties aan te praat te krijgen, hebben de client-side applicaties vergelijkbare problemen. Hoe meer originele onderdelen hergebruikt kunnen worden, hoe authentieker het eindresultaat is. Als we de grachtenpand analogie weer toepassen blijft de originele gevel intact, maar veel achter de gevel is opnieuw gebouwd. Of en hoeveel je moet restaureren hangt af van hoe compleet de gevonden data zijn.

In de praktijk kom je ook mengvormen tegen tussen emuleren en restaureren. Ook bij een emulatie zal het nodig zijn om wijzigingen te maken, zoals bijvoorbeeld het aanpassen van netwerkinstellingen en services en het veranderen van wachtwoorden.

Replicatie

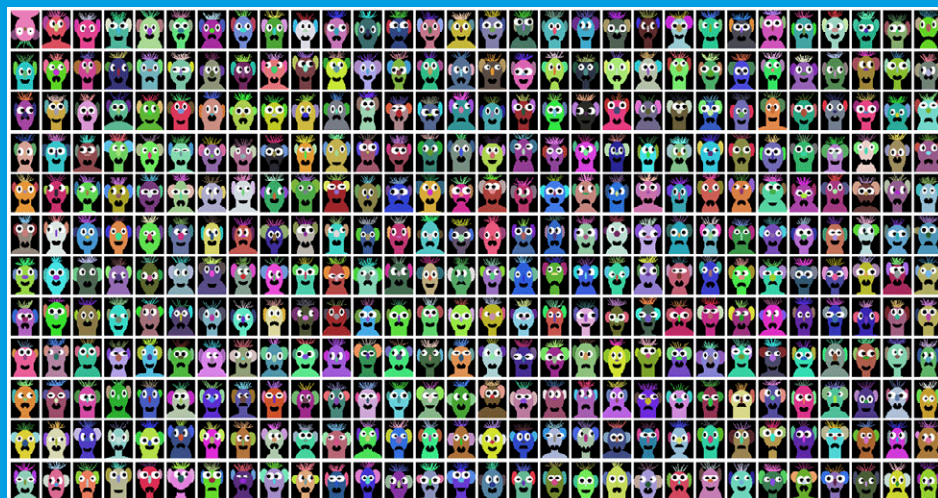
Een replicatie ziet er van buitenaf (de interface) hetzelfde uit als een emulatie of restauratie. Onder de motorkap is een replicatie echter compleet anders. Een replicatie bouwt vanaf de grond af aan de functionaliteit opnieuw op. Een replicatie zal waarschijnlijk gebruik maken van de originele data, maar zelfs dat is niet gegarandeerd. De data kunnen bijvoorbeeld in een ander formaat aanwezig zijn. Het

repliceren kan een aantrekkelijke optie zijn omdat zowel de emulatie als de restauratie de originele software gebruiken inclusief aanwezige veiligheidsrisico's. Ook kan in de replicatie rekening gehouden worden met privacykwesaties, copyright en andere wettelijke regelgeving. De replicatie is kortweg een veilige versie, geschikt voor publieke toegang. Een replicatie kan bijvoorbeeld onder een website van een museum gehangen worden, of in een museum getoond worden waar het publiek kan interacteren met het object. Een museum kan ook een replicatie op een USB-stick beschikbaar maken voor het publiek, of via een museumwinkel verkopen. Dit is niet mogelijk met een geëmuleerd of gerestaureerd object.

Onze aanpak: case study 'DDS Herleeft'

Een aantal studenten en verschillende oud-medewerkers van DDS zijn in de verschillende lagen klei (data) afgedaald en zijn aan de slag gegaan met de opgegraven data. Na bestudering van de data werd besloten om één van de unieke opgravingen als basis te nemen, te weten De FREEZE. Door de FREEZE te gebruiken betekent dat de bouwstenen voor de restauratiewerkzaamheden bestaan uit de 'momentopname' van 15 januari 1996. Om het digitale object, DDS, weer werkend te krijgen reconstrueerden de studenten uit De FREEZE twee versies van DDS3.0 (de pleinen- en huizen-interface): De Replica en De Emulatie. Maar eerst werd iets anders gevonden.

De Avatar-generator was het eerste resultaat van de restauratiewerkzaamheden. In 2015 groeven de studenten de Avatar-generator op en kregen deze weer werkend. De Avatar-generator stamt uit de vroege periode van DDS.

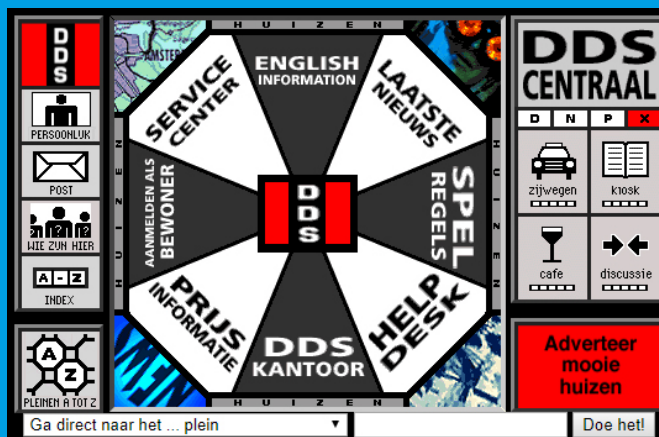


DDS Avatars, webarcheologische vondst

De avatars, ook wel dodo's genoemd (DoDoS, van DDS) deden hun intrede bij de opening van DDS3.0, de inmiddels bekende (stads)interface van DDS met de pleinen en huizen die op 10 juni 1995 gelanceerd werd. Elke nieuwe bewoner kreeg automatisch een avatar via een generator. De inspiratie voor de vormgeving van de avatars kwam uit de Muppet Show, van de Muppet Beaker, de laboratorium assistent (lees: proefkonijn) van Dr. Bunsen Honeydew.

De studenten interviewden de oorspronkelijke makers: Rob van der Haar (toen: interaction designer), Erwin Bolwidt (toen: systeembeheerder en programmeur), Michaël van Eeden (toen: programmeur en systeembeheerder) en Marjolijn Ruyg (toen: vormgever). Zij wilden het concept van een persoonlijke avatar introduceren voor 'wie zijn er op dit plein' en de webcafés. Na de overgang van de Bulletin Board System variant (DDS1.0) van DDS naar de web-variant (DDS3.0) was naar het gevoel van de makers en bewoners een 'vorm van plek' en 'aanwezig zijn' verloren gegaan. Om de bewoners een unieke identiteit te geven op een 'onpersoonlijk' internet ontwierpen de makers 'iets grafisch', de avatars. Ter herinnering: Slechts weinig mensen hadden in de vroege periode van DDS een digitale camera of tekenprogramma, zoals nu, om dit zelf te maken.

Het tweede resultaat van de restauratiewerkzaamheden van de studenten was De Replica. Deze Replica is een 'make-over' waarin de historische DDS-data met nieuwe technologie (moderne software) ontsloten worden. Naast het namaken van het historisch object waarbij de gebruikerservaring voorop stond, hebben de studenten de privacykwesties en beveiligingsproblemen aangepakt. Hierdoor is het mogelijk om deze versie in de toekomst in het museum en het e-depot te presenteren. De Replica omvat alleen die informatie die toegankelijk was zonder in te loggen op DDS. In De Replica zijn dus geen privacygevoelige data, zoals e-mail en accountgegevens, opgenomen. Door gebruik te maken van moderne software zijn ook de oude beveiligingsproblemen opgelost.



De Replica van DDS



Studenten laten De Replica (links) en De Emulatie (rechts) zien

Het derde resultaat van de restauratiewerkzaamheden was De Emulatie. De uitgangspunten voor De Emulatie waren het behoud van de gebruikerservaring en het herstel - zo authentiek mogelijk - van de originele software. De studenten hebben een deel van de functionaliteit weer aan de praat gekregen, met behoud van de complete originele structuur en de bestanden. De Emulatie bevat zowel de privacygevoelige data als de historische beveiligingsproblemen. De Emulatie is alleen voor wetenschappelijk onderzoek toegankelijk.

Technische verdieping: De Emulatie en afhankelijkheden

De FREEZE was een kopie van de drie servers waar de DDS op draaide. Dit was geen werkend systeem. De back-up bevatte voornamelijk data en geen OS. Het was dus niet mogelijk om het geheel onder een emulator te plaatsen. Dit maakte het noodzakelijk om alle componenten die wel gevonden waren op een nieuw systeem aan de praat te krijgen. Hiervoor werd gekozen voor een modern Linux-systeem. Om met de data weer een werkend systeem te maken moesten er eerst een aantal 'dependencies', afhankelijkheden, opgelost worden. DDS draaide immers op de Solaris SPARC-architectuur. Huidige computers draaien echter bijna altijd op Intel's x86 architectuur. Aangezien de source code aanwezig was bleek deze optie haalbaar.



Studenten nemen appeltaart in ontvangst (foto Daniël Ockeloën, CC BY-NC-SA 2.0)



Ontwerp studenten voor "De Digitale Stad gekraakt"

De studenten presenteerden hun vondsten in aanwezigheid van de DDS community (experts uit ons netwerk, oud-bewoners, oud-medewerkers en vrijwillige webarcheologen), onder de titel "De Digitale Stad gekraakt".⁸⁵

TIP: Houd je aan tradities!

DDS loofde een appeltaart uit aan iedereen die DDS kon kraken. DDS had deze traditie van het roemruchte hackersblad Hack-Tic overgenomen. Als hackers de zwakke plekken in een systeem aanwezen en ze de kwetsbaarheden niet misbruikten, beloonde Hack-Tic (en later de opvolger XS4ALL) de hackers met appeltaart. Tijdens ons project hebben we ons ruimschoots aan deze traditie gehouden en hebben we meerdere appeltaarten gebakken en uitgereikt aan studenten en vrijwillige webarcheologen die de opgegraven code weer aan de praat kregen.⁸⁶

⁸⁵ Bekijk de presentaties: <https://hart.amsterdam/DDSgekraakt>

⁸⁶ "Wij geloven dat hackers juist een heel belangrijke rol hebben bij het verbeteren van beveiliging: als een systeem zwakke plekken blijkt te hebben, dan is het in ieders belang dat die worden gefixt. En de beste manier om dat voor elkaar te krijgen is niet hackers te straffen, maar ze juist te belonen". (...) "Als jij een kwetsbaarheid ontdekt in een van de systemen van XS4ALL, dan willen we dat graag weten zodat we het probleem kunnen oplossen. Als jij belooft de kwetsbaarheid niet te misbruiken, dan beloven wij dat we geen aangifte doen. Je krijgt een heerlijke appeltaart als beloning en als we besluiten de gevonden kwetsbaarheid openbaar te maken, dan doen we dat in goed overleg met jou". "Hackers zijn geen boeven", door XS4ALL (27 augustus 2014), <https://blog.xs4all.nl/hackers-zijn-geen-boeven/>

Technische verdieping: De reconstructies geïllustreerd

Om de verschillen tussen de reconstructies van de studenten (De Emulatie en De Replica) te illustreren hebben we de volgende diagram opgenomen. In de linkerkolom zie je De Emulatie. Dit is een vorm van restauratie. Deze bevat alle originele FREEZE data en de originele webapplicaties. De Emulatie is volgens de theorie geen pure emulatie, want het draait niet op het originele OS (SUNOS), architectuur (SPARC) en originele webserver (Apache). Ook is de code opnieuw gecompileerd om deze weer werkend te krijgen op een nieuw Linux systeem. Het emuleert kortom wel de functionaliteit, met grotendeels de originele componenten. Daarom heet deze reconstructie De Emulatie. In de middelste kolom is het origineel van DDS weergegeven. Dit is exact zoals een emulatie eruit zou zien. Het enige verschil tussen het origineel en een emulatie is dat een emulatie draait op virtuele hardware. In de rechterkolom zie je De Replica. Deze bevat een deel van de originele FREEZE data, maar de webapplicatie is opnieuw gebouwd in PHP.

Om het mogelijk te maken om de verschillende reconstructies op hetzelfde systeem te gebruiken is er gebruik gemaakt van Linux-containers (LXC), zoals te zien is in de afbeelding. Dit is niet noodzakelijk voor de functionaliteit. Ook andere emulatie systemen zouden hiervoor gebruikt kunnen worden.

De Emulatie			De Replica		
Browser		De Digitale Stad		Browser	
NGINX		Browser		Apache	
code (Perl, C, Bash, HTML)	Freeze data	Apache		code (PHP, Bash, HTML, CSS)	Freeze data
compilers		code (Perl, C, Bash, HTML)	DDS data(in filesysteem)		
LXC		compilers		LXC	
	VM		VM		
linux		SUN OS		linux	
computer/server		computer/server		computer/server	

7.5 Duid de data

Er zijn verschillende methoden om informatie uit de oorspronkelijke bronnen te halen. Je kunt specifieke elementen onderzoeken, datasets bouwen en tools inzetten om de data doorzoekbaar te maken, te filteren en te visualiseren. Hierdoor kunnen toekomstige onderzoekers aan de slag met de unieke historische content; alhoewel hier nog (relatief veel) technische kennis voor nodig is.

Onderzoek specifieke elementen (alleen kijken)

Met het onderzoeken van een specifiek element bekijk je hoe jouw bron eruit ziet, hoe de structuur is en welke data de bron bevat. De techniek die hiervoor gebruikt wordt is al beschreven in 'Verken de data'. Bij dit onderdeel komen verder geen nieuwe technieken kijken. Het is niet altijd nodig het object te reconstrueren, bijvoorbeeld als het digitale object nog werkt of wanneer voor jouw doel geen werkend object nodig is. Het onderzoeken van de specifieke bestanden geeft vaak al veel inzicht in het object.

Onze aanpak: case study 'DDS Herleeft'

Tijdens het college History of Digital Cultures hebben meerdere studententeams deze aanpak gehanteerd. Zij verrichten thematische onderzoeken naar verschillende aspecten van De Digitale Stad. De studenten maakten gebruik van De FREEZE als bron, naast de externe bronnen. Een van de studentengroepen deed bijvoorbeeld onderzoek naar het gebruik van metaforen in DDS. Door bronnenonderzoek had ze al veel informatie gevonden. Ze gebruikte De FREEZE data om het gebruik van metaforen te staven en verder te verkennen. Een nieuwe tool of een werkende authentieke interface waren hiervoor niet nodig. De studenten konden door middel van de 'command line interface' van de Linux-werkomgeving de juiste bestanden vinden en openen.

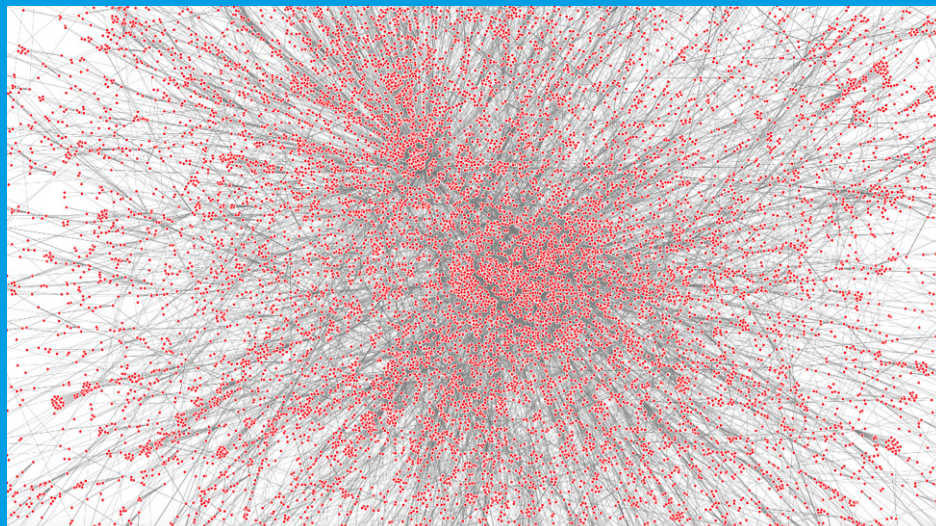
Herinterpretatie (oude data, moderne software)

In plaats van een complete reconstructie te maken is het ook mogelijk om de data te herinterpreteren. Je doet dit om een specifiek onderdeel te bestuderen, of om verder onderzoek te vergemakkelijken. Bij herinterpretatie doe je iets nieuws met het materiaal wat niet binnen het oorspronkelijke doel viel. Je steekt de data dus in een nieuw jasje (moderne software). Herinterpreteren is een handige methode om informatie uit het bronmateriaal makkelijker toegankelijk te maken.

Onze aanpak: case study 'DDS Herleeft'

Een aantal studenten heeft gebruik gemaakt van de herinterpretatie-methode. Ze hebben nieuwe software gemaakt om het bronmateriaal voor specifieke onderzoeksdoelen toegankelijk(er) te maken. Ze hebben datasets gemaakt en hebben de data gefilterd op bijvoorbeeld woorden, thema's, e-mails of gebruikersstatistieken.

Team 'Gabriele's Pizzeria'⁸⁷ deed bijvoorbeeld onderzoek naar de gemeenschappen in De Digitale Stad. Hiervoor maakten zij gebruik van de e-mails die meegenomen waren in De FREEZE. Door deze data te gebruiken en te visualiseren konden ze gemeenschappen visueel identificeren. Hieronder zie je een plaatje van het studententeam 'Gabriele's Pizzeria'. Je ziet het e-mailverkeer van de DDS zoals gevonden in De FREEZE. De rode stippen representeren e-mailadressen en de grijze strepen de e-mails.



Een ander studententeam, 'unfreeze', deed onderzoek naar de gebruikersstatistieken die te vinden waren in De FREEZE. Zo konden zij grafieken maken van het aantal posts op DDS door de tijd heen, wanneer gebruikers online waren en nog vele andere interessante feiten over hoe DDS gebruikt werd. Op de volgende pagina de factsheet van studententeam 'unfreeze', met de gebruikersstatistieken van DDS zoals gevonden in De FREEZE:

⁸⁷ Met dank aan de studententeams:

'Gabriele's Pizzeria': Kishan Nirghin, Tim Veenman, Millen Mortier, Robert Jansma, Randy de Vries, Ronald Bethlehem, Gabriele Di Bernardo.

'unfreeze': Alex van der Werff, Bastiaan Weijers, Boudewijn Braams, Onno Valkering, Robin van der Markt, Sam Ansmink.

Total number of users	50,462
Total number of chat messages	1,054,752
Total number of newsgroup messages	13,925
Total number of emails	313,994
Number of email users	23,018 (45% of total)
Average number of mails per user	14
Number of send mails	140,247
Number of received mails (direct, cc or bcc)	172,934
Number of emails by the most active email user	3,296
Users with less than 10 sent emails	12,317 (47%)
Number of chat users	8,993 (18% of total)
Average number of chats per person	117
Number of chats by the most active chatter	19,157
Users with less than 10 chat messages	5,276 (68% of chatters)
Number of newsgroup users	1,792 (4%)
Average number of newsgroup messages per person	8
Number of messages by the most active newsgroup user	599
Users with less than 10 newsgroup messages	1,593 (89%)
Participating users*	26,013 (52% of total)
Engaging users (at least 10 activities)	8,741 (17%)
Engaging users only active during weekdays (mo-fr)	3273 (37%)
Engaging users only active in the weekend	36 (less than 1%)
Users only active in chat	2,360 (5%)
Users only active in email	15,780 (31%)
Users only active in newsgroups	561 (1%)
Users active in all three categories	478 (less than 1%)

Gebruikersstatistieken van DDS zoals gevonden in De FREEZE

Technische verdieping: Structureer data in een database

Data die aanwezig zijn op een filesysteem, zoals het geval was bij De FREEZE data, zijn niet altijd makkelijk te gebruiken. De data zijn aanwezig, maar niet gestructureerd. Door eerst de gegevens van De FREEZE gestructureerd in een database te importeren konden de studenten vragen over DDS beantwoorden door middel van SQL-queries. De studenten importeerden verschillende data die vervolgens door het hele team gebruikt konden worden. Het importeren van een type data hoefde maar één keer gedaan te worden, waarna alle studenten de database konden gebruiken voor hun eigen analyses. In de twee bovengenoemde voorbeelden maakten de studenten dan ook gebruik van een SQL-database. Om je onderzoek te kunnen valideren moeten de authentieke data op het filesysteem wel beschikbaar zijn.

08

Ontsluiten

Hoe kun je je opgravingen duurzaam bewaren en toegankelijk maken en houden?

Je gaat nu aan de slag met het duurzaam archiveren van je digitale erfgoed en de langetermijntoegankelijkheid. Je hebt in de vorige hoofdstukken kunnen lezen over de methoden waarmee de data en dynamische systemen opgegraven en gereconstrueerd kunnen worden. Het bewaren van een dynamische webomgeving vraagt andere methoden dan van een (semi-)statische webomgeving. Bij dynamische webomgevingen focus je op de dynamiek en interactiviteit van je digitale object.

Hier komt een en ander om de hoek kijken. Websites zijn gebaseerd op technologie die aan snelle verandering onderhevig zijn. Denk bijvoorbeeld aan de webbrowser die de site weergeeft, het besturingssysteem waar die browser in opereert, de bestandstypes en programmeertalen die zijn gebruikt om de website op te bouwen en de computeronderdelen waar dit alles op draait. Hoe kunnen data en dynamische systemen opgenomen worden in een digitaal archief? Welke voorbereidingen zijn nodig om duurzame archivering mogelijk te maken? Ons onderzoek voor onze case study is nog volop in ontwikkeling. In dit hoofdstuk beschrijven wij onze wensen en voorbereidingen. Om zelf aan de slag te kunnen gaan bieden we een stappenplan en praktische handvaten. Je gaat kijken naar *wat* je archiveert, *hoe* je het archiveert, *voor wie* je het archiveert, *welke afspraken* je moet maken en *hoe* je toegang geeft tot het archief.

Stappenplan:

- 1 Archiveer: statisch of dynamisch?
- 2 Zoek een partner
- 3 Beschrijf de gewenste preservingsscenario's
- 4 Geef toegang aan doelgroepen
- 5 Documenteer, documenteer, documenteer

8.1 Archiveer: statisch of dynamisch?

Als je te maken hebt met een eenvoudige statische website kun je in principe zelf aan de slag. We noemen hier twee - vaak complementaire - methoden.⁸⁸ In de meest simpele variant sla je de broncode (platte HTML-pagina's) en je database op en je zet deze bijvoorbeeld op een harde schijf. In een meer geavanceerde variant archiveer je 'snapshots'. Je hebt hiervoor een 'webharvester' of 'webcrawler' nodig. Een Open Source voorbeeld hiervan is Heritrix.⁸⁹ De software maakt een lokale kopie van een website door webpagina's en bijbehorende grafische elementen en documenten naar een harde schijf te kopiëren. Zo wordt van een website een statische momentopname gemaakt die bewaard kan worden als een 'WARC' (Web ARChive). Als je zelf niet over de software en een infrastructuur beschikt kan je met een e-depot samenwerken.⁹⁰ Je kunt ook zelf webpagina's toevoegen aan de Wayback Machine of bij Archive.is.⁹¹ Let wel op: In dit geval bewaar je één specifieke pagina op één moment, niet meerdere pagina's, directories of hele sites. Dit is zeer bewerkelijk en biedt geen oplossing voor de langetermijn.

Als je een dynamische webomgeving wilt archiveren, voldoet het maken van momentopnames ('snapshots') niet en moet je een stap verder gaan. Bij dynamische webomgevingen focus je op de dynamiek en interactiviteit van je digitale object. Bij dynamische webarchivering bied je toegang tot (weer) werkende digitale objecten. Door de webobjecten duurzaam te archiveren houd je digitaal erfgoed toegankelijk voor toekomstige generaties. Deze nieuwe variant van preserving, het duurzaam en dynamisch archiveren, vergt de ontwikkeling van nieuwe technologie en diensten. Zowel nationaal als internationaal wordt gezocht naar technische oplossingen, maar er bestaat voor deze nieuwe variant nog geen methodiek die afdoende is.

TIP: Neem je website op met Webrecorder

Bekijk de (gratis) tool voor het deels dynamisch archiveren van websites van Rhizome.⁹² Webrecorder is een nieuwe open source tool voor het maken van (beperkte) dynamische, dus werkende en afspeelbare, webarchieven. Voor een complex systeem als DDS met meerdere (historische) applicaties voldoet de tool (nog) niet. Voor blogs met 'embedded' onderdelen en social media is deze tool goed te gebruiken.

⁸⁸ "Websites archiveren in 4 stappen" (2007), edavid.be. Zie: http://www.edavid.be/digitaaldepotproject/documenten/folders/pdf/folder3_online.pdf

⁸⁹ Zie <http://crawler.archive.org/>

⁹⁰ Bekijk de activiteiten van en de mogelijkheden bij de Expertgroep Webarchivering <http://www.ncdd.nl/kennis-en-advies/expertgroepen/expertgroep-webarchivering/>.

⁹¹ Ga naar <https://archive.org/web> (zie de "Save Page Now") of naar <http://archive.is/>.

⁹² <https://webrecorder.io/>.

Onze aanpak: case study 'DDS Herleeft'

In ons project archiveren we de originele en bewerkte DDS-data. Die bestaan uit de images van De FREEZE tapes en de data van de SUN SPARC-servers en het audiovisueel materiaal (primaire bronnen). Daarnaast bewaren we ook de reconstructies van DDS 3.0, De Replica en De Emulatie, de datasets van de studenten (secundaire bronnen) en de overige verzamelde data op de 'Grave Diggers Party', verslaggeving en publicaties over DDS, projectdata en onderzoekspapers van de studenten (overige bronnen). Ons doel is het bewaren van dynamische data in een dynamische webomgeving.

8.2 Zoek een partner

In Nederland zijn er verschillende nationale en regionale instellingen bezig met het archiveren van websites; bibliotheken, archieven, wetenschappelijke organisaties en musea. Elke instelling hanteert eigen selectiecriteria en heeft verschillende middelen om die doelen te bereiken. Op nationaal niveau wordt samengewerkt aan een netwerk van landelijke voorzieningen om digitale informatie duurzaam te behouden. Dit netwerk wordt door het Netwerk Digitaal Erfgoed (NDE) en de Nationale Coalitie Digitale Duurzaamheid (NCDD) gecoördineerd. Het netwerk heeft in kaart gebracht wat nodig is en wat nog ontbreekt, en biedt een overzicht van de voorzieningen, diensten en kennis die in Nederland beschikbaar zijn.

TIP: Overzicht webarchivering en webwijzer duurzame voorzieningen

Als museum of collectiehoudende instelling heb je voorzieningen, diensten en kennis nodig om digitale materialen te archiveren. Bekijk wat je zelf kan en wat je moet uitbesteden. Als je onderdelen mist kan je in Nederland op verschillende plaatsen samenwerking zoeken. Verschillende aangesloten instellingen bieden verschillende voorzieningen, diensten en kennis. Zoek jij een webarchief of wil je meer weten over wat er in Nederland gebeurt op het vlak van webarchivering, kijk dan bij het Kennisplatform webarchivering.⁹³ Het Netwerk Digitaal Erfgoed (NDE) en de Nationale Coalitie Digitale Duurzaamheid (NCDD) hebben in kaart gebracht wat nodig is en wat nog ontbreekt voor een goede landelijke infrastructuur voor het duurzaam behoud van het digitaal erfgoed. Bekijk bijvoorbeeld de 'Voorzieningenvinder Digitale Duurzaamheid'.⁹⁴

⁹³ <https://informatie2020.pleio.nl/groups/profile/48637242/kennisplatform-webarchivering>

⁹⁴ Zie: <http://voorzieningen.ncdd.nl/> (bètaversie). Lees ook: 'Samen bouwen aan een netwerk van landelijke voorzieningen. Eindrapport van het onderzoek naar een landelijke infrastructuur voor duurzame toegang tot digitale informatie, (2015) door Joost van der Nat en Marcel Ras, http://ncdd.nl/site/wp-content/uploads/2015/09/NCDD-onderzoek_2015_DEFWEB.pdf

Onze aanpak: case study 'DDS Herleeft'

Het Nederlands Instituut voor Beeld en Geluid, partner in ons project, heeft duurzame digitale preservering bestempeld tot primaire taak. In 2016 heeft Beeld en Geluid het Data Seal of Approval⁹⁵ (DSA) voor Trustworthy Digital Repositories (TDR) ontvangen omdat het complexe digitale archiveringsproces aantoonbaar beheerst. Het domein van de digitale preservering⁹⁶ strekt zich uit over de complete levenscyclus van digitale objecten; dat wil zeggen over alle processen rond de instroom, de opslag, het beheer en het verlenen van toegang tot dat object. Beeld en Geluid heeft gekozen om haar webarchiveringswerkzaamheden te verbreden. Naast het archiveren van websites van publieke omroepen voert het instituut pilots uit om de duurzame opslag van en toegang tot nieuwe media, games en internetcultuur te onderzoeken. Hierbinnen is de ruimte ontstaan om dynamische webcontent op meer experimentele wijze te archiveren en toegankelijk te houden. In dit kader heeft Beeld en Geluid besloten om de werkende versie van DDS3.0, en overige opgegraven DDS data, te archiveren. Samen met Beeld en Geluid zijn we uitgekomen op de volgende preserveringsscenario's.

8.3 Beschrijf de gewenste preserveringsscenario's

Onze aanpak: case study 'DDS Herleeft'

We kunnen niet genoeg benadrukken hoe bijzonder en belangrijk het werk is geweest dat DDS heeft verricht. In de tijd van DDS was er nog lang geen sprake van archiveer- of verzamelbeleid voor nieuwe media objecten. Dankzij de bijzondere toekomstvisie van de jubilerende DDS-ers en hun tijdcapsule uit 1996 zouden wij niet zover gekomen zijn als we nu zijn.

⁹⁵ Zie: <https://www.datasealofapproval.org>

⁹⁶ "Onder het begrip digitale preservering wordt binnen het beleidskader van Beeld en Geluid het volgende verstaan: Het geheel aan activiteiten en processen dat noodzakelijk is voor de intellectuele en technische instandhouding van de digitale collecties door de tijd heen, met als doel de collecties duurzaam toegankelijk te houden voor de gebruikersgroepen. Hoofddoel van digitale preservering is het behoud van integriteit en authenticiteit van de digitale objecten." Zie: "Digitale Preservering Beeld en Geluid: Beleid, Standaarden en Procedures. Netherlands Institute for Sound and Vision." (2016), Annemieke de Jong. <http://publications.beeldengeluid.nl/pub/387>, pagina 19. En: "Naar een multimediale toekomst in Beeld en Geluid: Beleidsplan 2016-2020" (2015), Nederlands Instituut voor Beeld en Geluid. <http://files.beeldengeluid.nl/beleidsplan/2016-2020/>.

In de voetsporen van De Digitale Stad willen we ons archief beschikbaar stellen “ter bestudering voor archeologen in een verre toekomst”. Wij hebben ruim acht miljoen bestanden teruggevonden. Een aantal onderdelen van DDS hebben we weer werkend gekregen. Onze vraag is nu hoe we dit materiaal duurzaam (en deels ook het liefst werkend) bij elkaar kunnen houden en kunnen archiveren. We hebben voor drie preserveringsscenario's gekozen, omdat deze drie de duurzame opslag van en toegang waarborgen. Op dit moment zijn wij nog volop in het onderzoek naar de scenario's samen met Beeld en Geluid. We kunnen dus nog geen uitkomsten opnemen in het handboek, maar wel onze gewenste preserveringsscenario's.

Drie preserveringsscenario's:

1. **Bit-preserved** (niet openbaar toegankelijk, read only). Alleen opslag van de data (de bits en bytes). Bit-preserved is vergelijkbaar met De FREEZE van De Digitale Stad. Je bewaart alle data die je tijdens je project hebt opgegraven, zowel in originele als bewerkte staat. Deze bestanden bestaan in ons geval uit: De originele DDS-data: de images van De FREEZE tapes en de data van de SUN SPARC servers en het audiovisueel materiaal (primaire bronnen), de reconstructies van DDS 3.0, De Replica en De Emulatie, de datasets van de studenten (secundaire bronnen) en de overige verzamelde data op de 'Grave Diggers Party', verslaggeving en publicaties over DDS, project data en onderzoekspapers van de studenten (overige bronnen). Het gaat hier om het bewaren van de enen en nullen, zonder duurzaamheidsacties. Deze back-up kan in de toekomst worden gebruikt om een meer complete reconstructie te maken.
2. **Duurzame en dynamische opslag** (niet openbaar toegankelijk, read only). Op dit moment zoeken we een methode waarmee we de uitgebreide en diverse DDS-collecties met dynamische (werkende) data duurzaam kunnen bewaren. We onderzoeken de mogelijkheid om de reconstructies (De Replica en De Emulatie) en de back-up van alle data op te kunnen slaan als virtuele machines, en deze als zodanig onder te brengen en toegankelijk te maken. Om te garanderen dat de data in de toekomst ook blijven functioneren moet het gearchiveerde digitale object met regelmaat gemonitord worden, zodat risico's en veranderingen die gevolgen kunnen hebben voor de digitale collectie tijdig worden opgemerkt, zodat de juiste maatregelen getroffen kunnen worden (preservation planning). Zowel de technologie als de dienstverlening voor deze nieuwe methode van dynamische preserved zijn in ontwikkeling bij Beeld en Geluid.
3. **Webarchiveren** (openbaar toegankelijk, read only). Opnemen van Web ARChive (WARC) bestanden van De Replica. WARC is een veelgebruikt archiveringsformaat voor het opslaan van webcrawls.

Onze aanpak: case study 'DDS Herleeft'

Technische verdieping: De reconstructies als virtuele machines

Er is voor gekozen om de twee reconstructies, De Emulatie en De Replica, op te slaan als virtuele machines. Een virtuele machine is, zoals de naam suggereert, een machine maar dan virtueel. Neem De Replica als voorbeeld. De Replica draait de software die nodig is om DDS te draaien. Deze software werkt alleen op een Linux machine met een x86 architectuur. Als je de software naar een andere machine wilt overzetten is het nodig de source code te hercompileren, de nodige programma's voor de nieuwe machine te installeren en alle andere afhankelijkheden weer opnieuw op te lossen. Virtuele machines lossen dit probleem voor je op. Zo kan de virtuele machine van De Replica op een machine worden gezet met een ander operating systeem. Ook een andere computer architectuur is geen probleem, zolang deze maar software heeft die deze virtuele machine ondersteunt. Een nieuw systeem hoeft niet meer alle software van De Replica te ondersteunen maar hoeft alleen in staat te zijn om de virtuele machine te draaien. Door een virtuele machine te gebruiken reduceer je het aantal afhankelijkheden tot slechts één. Virtuele machines zijn daarom ideaal als archiveringsformaat voor dynamische digitale objecten.

TIP: Gebruik het OAIS-model⁹⁷

Er zijn een aantal processen waar je rekening mee moet houden om digitaal erfgoed duurzaam te bewaren en te ontsluiten in een e-depot. Zowel de aanbieder van het digitale object als de ontvanger (het e-depot) zijn hierbij betrokken. Je kunt hiervoor het 'Open Archive Information System'- model (OAIS) gebruiken.

We hebben het OAIS-model als leidraad gebruikt in ons project. We hebben per proces onze vragen toegevoegd:

- **Ingest:** Het binnenhalen van de gegevens (digitale objecten en metadata). Hoe worden DDS data ingelezen in het archief en welke afspraken kunnen worden gemaakt?
- **Archival storage:** Het opslaan van de digitale objecten. Op welke wijze dienen de DDS-data opgeslagen te worden?
- **Datamanagement:** Het beheer van de metadata van de digitale objecten ten behoeve van zoek-, opvraag- en beheerfuncties en het uitvoeren van controles. Hoe dienen de DDS data beheerd te worden en welke metadata zijn er nodig hiervoor?

⁹⁷ Lees: "How can we improve our web collection? An evaluation of webarchiving at the KB National Library of the Netherlands (2007–2017)", door: Barbara Sierman en Kees Teszelsky (29 august 2017), <http://journals.sagepub.com/doi/10.1177/0955749017725930>, "Duurzaamheidsbeleid" door Barbara Sierman en Marcel Ras, OD (mei-juni 2017), <http://www.ncdd.nl/news/artikel-duurzaamheidsbeleid-od-magazine/> en "Het OAIS-model, een leidraad voor duurzame toegankelijkheid", (2012) door Barbara Sierman, in: "Handboek Informatiewetenschap", <http://www.iwabase.nl>.

- **Administration:** Het beheer van het e-depot: het coördineren van de activiteiten (diensten, services en functionaliteiten).
- **Preservation planning:** Het monitoren van het archief. Het plannen van het duurzaam beheer van de digitale objecten. Hoe blijven de DDS-data in de toekomst toegankelijk en hoe organiseren we het beheer hiervan? Dit is de kern van het duurzaam opslaan en toegankelijk houden voor toekomstige generaties van je digitale object.
- **Access:** Toegang geven tot digitale objecten of informatie aan gebruikers. Hoe kunnen geïnteresseerden de DDS-data bereiken?

8.4 Geef toegang aan doelgroepen

In de preserveringsscenario's leg je ook vast wie de beoogde en toekomstige gebruikers zijn van je gearchiveerde digitale objecten.

Onze aanpak: case study 'DDS Herleeft'

In ons project hebben wij twee doelgroepen voor ogen: Het publiek en onderzoekers. Zoals beschreven in het vorige hoofdstuk hebben de studenten om die reden twee reconstructies ontwikkeld, De Replica en De Emulatie. In De Replica (oude data in nieuwe software) zijn de privacykwesties en beveiligingsproblemen aangepakt. De studenten hebben de privacygevoelige data niet meegenomen. Deze variant kan lokaal in een museum getoond worden. De Emulatie (oude code op nieuwe machines) bevat wel de privacygevoelige data en heeft historische beveiligingsproblemen. De Emulatie en de overige gevonden data worden alleen voor wetenschappelijk onderzoek toegankelijk gemaakt en zijn alleen in het e-depot op te vragen. Als onderzoekers De Emulatie en de overige gevonden data willen bekijken, kunnen ze in de toekomst een verzoek indienen bij de stichting Beheer Digitaal Erfgoed DDS bestaande uit de huidige projectpartners, die waakt over het (her)gebruik van de DDS-data. De stichting beoordeelt de aanvragen van de onderzoekers.

8.5 Documenteer, documenteer, documenteer

We kunnen het niet vaak genoeg benadrukken: Voor het overdragen van je digitale object is het vastleggen van alle stappen in het webarcheologische proces (plannen, graven, reconstrueren en ontsluiten) en de overdracht van je digitale erfgoed aan een e-depot van groot belang. Documenteer, documenteer, documenteer! Goede documentatie is de sleutel voor nieuwe (her)interpretaties en (re)constructies door toekomstige webarcheologen, (her)gebruik door onderzoekers en het duurzaam behoud van ons digitale erfgoed. En daarmee houd je het verleden levend voor het publiek in de toekomst.

09

Let the Bytes Free!

In dit DIY (Do It Yourself) Handboek doen we verslag van een webarcheologische speurtocht naar nieuwe webobjecten, met nieuwe instrumenten en nieuwe onderzoeksvragen. De processen, plannen, graven, reconstrueren en ontsluiten, hebben we stap-voor-stap beschreven. In helikopterperspectief belichten we nu onze ‘lessons learned’, vatten we onze resultaten en bevindingen uit de case study ‘DDS Herleeft’ samen, geven we een aanzet tot toekomstig onderzoek en doen we twee oproepen.

Onze ‘lessons learned’ zijn:

- Identificeer belangrijk en kwetsbaar digitaal erfgoed en grijp (op tijd) in. Als we niets doen dreigt een deel van ons weberfgoed verloren te gaan.
- Zet de kracht in van community building en crowdsourcing. Ga schatgraven naar (verloren) digitaal erfgoed. Organiseer bijeenkomsten en breng mensen, verhalen en objecten bij elkaar. Breng je team samen met de oorspronkelijke bouwers, bedenkers en bezoekers van je digitale object. De eerste ontwikkelaars zijn een bron van kennis en ervaring en kunnen uitleg geven over de historische context van het object en bieden reflectie op andere tijden en ontwikkelmethoden.
- Het beheer en behoud van digitaal erfgoed begint bij de creatie. Ondersteun waar mogelijk de makers van het nieuwe digitale erfgoed. Denk bij acquisitie of overname van born digital materiaal samen na over het toekomstig behoud van born digital objecten, bijvoorbeeld over het gebruik van software, bestandsnamen, opslag en back-up, juridische en ethische aspecten en bespreek de mogelijke duurzaamheidsstrategieën en zorg voor goede documentatie over de werking, het gebruik en de betekenis.
- Bevorder het (her)gebruik van je digitale erfgoed. Maak de digitale objecten vindbaar, toegankelijk, interoperabel, herbruikbaar en sla ze duurzaam op. (Her)gebruik van collecties door een grote gebruikersgroep bevordert het belang en de noodzaak van hun behoud.
- Documenteer en deel je kennis. Neem naast tekstuele beschrijvingen ook foto's, screenshots, screencasts en video's op. Documentatie is de sleutel voor (her)gebruik en het levend houden van het weberfgoed. Door het delen verlaag je drempels voor andere toekomstige webarcheologische projecten.
- Nieuwe objecten vergen nieuwe instrumenten en nieuwe onderzoeksvragen. Door de snelle verandering van techniek moet je je voorbereiden op nieuwe werkwijzen en samenwerkingsverbanden. Zoek de samenwerking op met nieuwe partners uit verschillende disciplines.

Onze aanpak: case study 'DDS Herleeft'

Onze case study 'De Digitale Stad Herleeft' loopt als rode draad door het handboek. Wij stelden ons de volgende vragen: Hoe graaf je De Digitale Stad op en transformeer je het van een virtueel Atlantis in een virtueel Pompeï? Hoe reconstrueer, bewaar, ontsluit en presenteer je DDS materiaal op een duurzame manier voor toekomstige generaties? We zijn al flink gevorderd, en we zijn er bijna. We begeven ons continu in een spanningsveld tussen de noodzaak om direct te handelen, en de (tijdrovende) zoektocht naar middelen en de ontwikkeling van nieuwe methoden. We hebben (nog) geen compleet opgegraven stad. We bieden geen kant-en-klare methode en geen complete webarcheologische gereedschapskist. Wat we wel bieden is een eerste aanzet tot een mooi vakgebied, webarcheologie, en de aanzet tot een nieuwe manier van dynamisch archiveren van webobjecten.

Onze resultaten en bevindingen:

- **Afbakening van het object: data.** We hebben (tot nu toe) brokstukken uitgegraven. Na bestudering hebben we besloten om de historische data uit de unieke opgraving van de FREEZE, als basis te nemen. Dit betekent dat onze bouwstenen uit een 'momentopname' komen van 15 januari 1996. Uit de data hebben we (tot nu toe) alleen de derde variant van de stad (DDS3.0) opgegraven en gereconstrueerd. Dankzij gesprekken met oud-bewoners en oud-medewerkers weten we inmiddels dat na de overstap naar de grafische interface (DDS2.0 en 3.0) de eerste telnet/command-line versie (DDS1.0) ook nog draaide. Je kunt dus niet zeggen dat DDS1.0 tekst-gebaseerd was en DDS 3.0 grafisch. DDS was het geheel aan omgevingen. Daarnaast was DDS niet alleen een digitale omgeving. DDS bestond meer dan uit de virtuele stad alleen, DDS staat ook voor de vele projecten die DDS uitvoerde. Deze projecten vallen buiten onze scope. We hebben dus niet de complete stad met alle stadsgezichten over alle perioden uitgegraven en herbouwd.
- **Ontsluiting van het object: twee reconstructies.** De studenten hebben twee versies van de derde interface (DDS3.0) gereconstrueerd die we hopelijk in de toekomst kunnen ontsluiten. Ten eerste De Replica voor publiek, alleen in museum en e-depot toegankelijk. Ten tweede De Emulatie, alleen voor wetenschappelijk onderzoek en in het e-depot toegankelijk. Onderzocht is hoe 'authentiek' (historisch getrouw) het oorspronkelijke digitale object DDS weer tot leven gebracht kon worden. Daarnaast speelden juridische en ethische kwesties een rol. Privacy is altijd een belangrijk issue geweest in DDS. Wij onderschrijven dit en doen onze uiterste best om deze blijvend te waarborgen. De publieke variant, De Replica, omvat daarom alleen die informatie die toegankelijk was zonder in te loggen (voor ingewijden: de versie van de stad die je kon zien als je als 'toerist') en zijn de privacygevoelige en niet-openbare data, zoals accountgegevens en e-mails, niet opgenomen. Door gebruik te maken van moderne software zijn ook de beveiligingsproblemen aangepakt. De wetenschappelijke variant, De Emulatie, omvat wel de privacygevoelige data en heeft verouderde (historische) beveiligingsproblemen.

- **Ontsluiting van het object: geen exacte weergave.** Beide reconstructies zijn niet compleet. De beide varianten bestaan uit de pleinen en uit de 1.348 huizen. Er ontbreken nog webpagina's (zoals 'Verzamelgebouwen'), applicaties (zoals nieuwsgroepen en cafés), dynamische elementen (zoals audio, video en forums) en twee interfaces (DDS1.0 en DDS2.0). De studenten hebben de 'scherven van de gebroken vaas' zo goed als het kon weer aan elkaar geplakt, maar er ontbreken nog stukjes. Om de stad zo dicht mogelijk bij de oorspronkelijke versie te krijgen moeten we restauratiewerkzaamheden verrichten en meer onontgonnen data onderzoeken en werkend maken. En om de gebruikerservaring nog dicht bij de beginjaren van het web te brengen zou je echter ook de browserweergave, netwerkverbindingssnelheid, monitorresolutie en het computerbesturingssysteem moeten nabootsen. Vraag of is natuurlijk of de huidige generatie de traagheid van toen wel wíl ervaren.
- **Archivering van het object: klaar en nog in ontwikkeling.** De bit-preservering is inmiddels gelukt. De primaire digitale bronnen (de originele DDS data, de images van De FREEZE-tapes en SUN SPARC-servers en het audiovisueel materiaal), de secundaire digitale bronnen (de reconstructies van DDS 3.0, De Replica en De Emulatie, en de datasets van de studenten, en de overige data (de context, de verzamelde data op de 'Grave Diggers Party', verslaggeving en publicaties over DDS, project data en onderzoekspapers van de studenten) zijn gearchiveerd en bevinden zich op een draagbare harde schijf bij Beeld en Geluid. De duurzame en dynamische opslag van de DDS-data als virtuele machine bestanden zal in de loop van 2018 uitgevoerd worden bij Beeld en Geluid. De toegang in een museum en het e-depot volgen in het jaar 2019.
- **Utopie.** In onze wilde beginfase droomden we over het creëren van een representatieve variant van het internet-historisch monument DDS in een '4D-collectie' (lengte, breedte, diepte én tijd). We wilden de bezoekers door de tijd heen door de drie verschillende versies van De Digitale Stad laten wandelen en zo de evolutie van deze unieke virtuele stad en het vroege internet laten ervaren. Utopisch, en (nog) niet haalbaar. Ofschoon er onder de opgegraven 'klei' nog vele verscholen schatten op ons liggen te wachten is de vraag of we alle benodigde data hebben. Een andere vraag is of we alle (rauwe) data aan de praat krijgen. Daarnaast was DDS een complex systeem, ge(ver/her)bouwd door de jaren heen. Er zijn dus vele kwetsbare afhankelijkheden op het niveau van hardware en software. Je hebt bijvoorbeeld verschillende browser software, markup talen, programmeertalen, scripttalen, besturingssystemen etc. Wellicht kan de volgende generatie toekomstige webarcheologen een '4D-collectie' creëren?
- **Iteratief proces.** Het proces dat we tot nu gevolgd hebben is geen lineair proces geweest. Webarcheologie is een -bij uitstek- iteratief proces (vooral bij de technische digitale archeologie). We voeren kleine projecten uit, die elk een eigen te doorlopen workflow hebben. Waar

het kan bouwen we iedere keer weer voort op het werk dat al gedaan is. Neem bijvoorbeeld de reconstructies. Veel van de projecten gingen in eerste instantie niet veel verder dan het verkennen van beperkte delen van de opgegraven data. Deze 'herinterpretaties' leverden weer bouwstenen op voor een volgende fase of projectonderdeel.

- **Noodzaak tot samenwerking.** De huidige generatie erfgoedprofessionals en computer wetenschappers staan nog ver uit elkaar. Ze spreken elkaars taal (letterlijk) niet en de vakgebieden liggen nog mijlenver uit elkaar. Je kunt een historicus niet de (rauwe) DDS-data geven, het is voor hem of haar een onontwarbare abstracte kluwen nullen en enen. En computer wetenschappers kunnen doorgaans geen historische afstand nemen en de gevonden digitale objecten in de context van de tijd zien: Alsof een tijdmachine de code heeft verknoeid.⁹⁸ Hoe sla je nu een brug tussen beiden? Het goede nieuws is dat in de laatste jaren historici en informatici meer en meer vrienden zijn geworden waarvan het ontstaan van het nieuwe vakgebied *digital humanities*⁹⁹ getuigt (en ons project!). De digitalisering van vele bronnen heeft nieuwe mogelijkheden voor historisch onderzoek geopend. Informatici helpen de historici met het vindbaar, toegankelijk, uitwisselbaar en herbruikbaar maken van de data. Zijn er tech-savvy erfgoedprofessionals informatici nodig of geschiedkundige informatici? Het punt is dat je de kennis en de context van het digitale object nodig hebt, zowel op het vlak van sociale digitale archeologie als technische digitale archeologie, en dat je een gereedschapskist met instrumenten moet kunnen bedienen. In ons project kwamen de verschillende disciplines samen door de samenwerking tussen de projectpartners. Door deze samenwerking hebben we al veel bereikt en blijven we elkaar uitdagen.
- **Co-creatie en participatie.** De Digitale Stad is met vele handen gebouwd, en wordt nu met vele handen opgegraven, gereconstrueerd en bewaard voor de toekomst. Zonder de bijdragen (in de vorm van kennis, donaties, verhalen, subsidie, aandacht) van onze projectpartners, ervaringsdeskundigen (de oorspronkelijke bewoners van de stad: de bouwers, bedenkers, bezoekers en bewoners van DDS), studenten, vrijwilligers en experts uit ons netwerk waren wij niet waar we nu zijn. Wij zijn de deelnemers enorm dankbaar. Samenwerking maakt veel zaken mogelijk die anders buiten ons bereik zouden liggen. Het heeft ons inspiratie gebracht en heeft onze horizon verbreed.

⁹⁸ Twee voorbeelden van het leren denken met historische afstand die we geleerd hebben tijdens overleggen met de computer science studenten. Als een student zegt: "Dat is slecht geprogrammeerd", antwoord je: "Het is een interessante ontwerpkeuze". Als een student zegt: "Hoe konden ze dat nu denken?", zeg je: "Ze dachten anders in die tijd. Binnen de mogelijkheden van toen maakten zij die keuzes. We zijn nu (in het geval van DDS) drieëntwintig jaar verder en er is veel doorontwikkeld en veranderd".

⁹⁹ "*Digital humanities* verbindt methoden uit de klassieke humaniora met die uit de informatica om zo grote databestanden te onderzoeken en interpreteren. En door die nieuwe beschikbaarheid van data en instrumenten kunnen we de samenhang tussen tekst, beeld, en context onderzoeken". Bron: <https://blog.nwo.nl/humanities/digital-humanities/> "Digital Humanities" door José van Dijck (25 feb 2014)



DDS in vaste opstelling van het Amsterdam Museum

9.1 Toekomstig onderzoek

- Onder de klei liggen nog vele schatten verscholen. De onontgonnen data vormen een rijke bron aan geschiedenis voor toekomstige webarcheologen. We hebben nog een aantal archieven van de 'Grave Diggers Party' die nog niet bestudeerd zijn. Denk ook bijvoorbeeld aan data in webarcheologische depots, zoals bij de Koninklijke Bibliotheek of het Internationaal Instituut voor Sociale Geschiedenis.¹⁰⁰ Hier zijn data opgeslagen, maar nog niet uitgewerkt. Door de data te gebruiken blijven ze levend.
- De reconstructies van DDS3.0 (De Replica en De Emulatie) zijn nog niet af. Er zijn restauratiewerkzaamheden ('restorative curation') nodig. Dit is een nieuw gebied. Doel hiervan is om de stad zo dicht mogelijk bij de oorspronkelijke versie te krijgen.
- Net zoals bij archeologie zijn er meerdere lagen te bestuderen, naast de kleilagen (software en hardware) heb je cultuurlagen (data). Als straks de DDS data ontsloten worden is het tijd voor interpretatie en duiding door de toekomstige onderzoekers zoals geesteswetenschappers en mediahistorici.
- Dit handboek richt zich vooral op de erfgoedprofessionals. Het is een eerste aanzet. Wat nodig is, is een technisch DIY handboek geschreven voor de technische webarcheologen. En een voor historici...
- Naast wensen voor toekomstig onderzoek doen we hier ook een oproep: Wie start en onderhoudt de twee voor webarcheologie ontbrekende essentiële (e-)depots? Ten eerste missen we een centraal depot voor tools om verschillende media uit te kunnen lezen, zoals bijvoorbeeld tape readers. Ten tweede is er behoefte aan een softwarebibliotheek. Wie archiveert de oude software, naast het NIST in de Verenigde Staten, zoals oa operating systems (Solaris, Windows, MacOS), webserversoftware (Apache, NGINX), Office Suites (Microsoft Office, OpenOffice, Libre Office) etc.?

9.2 Deel jouw onderzoek

Met welk webarcheologisch onderzoek ben jij bezig? Deel jouw kennis en ervaringen met ons. Mail¹⁰¹ ons!

¹⁰⁰ Zie bij de Koninklijke Bibliotheek, <https://www.kb.nl/organisatie/onderzoek-expertise/e-depot-duurzame-opslag/webarchivering> of het Occasio Digital Social History Archive van het Internationaal Instituut voor Sociale Geschiedenis, <https://socialhistory.org/nl/collections/occasio>.

¹⁰¹ Tjarda de Haan: info@bitsandbytesunited.com

10

Bijlage I: Bronnen

Dit is geen uitputtende lijst, maar een lijst met publicaties en bronnen die ons project enorm geholpen en geïnspireerd heeft.

Publicaties:

- Inge Angevaere, Nationale Coalitie Digitale Duurzaamheid (NCDD). Born-digital: (hoe) bewaar je dat? <http://www.ncdd.nl/documents/20091130Born-digital.pdf>.
- Niels Brügger, Ralph Schroeder (editors), The Web as History: Using Web Archives to Understand the Past and the Present. UCL Press, 2017.
- Janet Delve, David Anderson (editors), Preserving Complex Digital Objects. Facet Publishing, 2014.
- DEN, Thema Digitale duurzaamheid, <https://digitalecultuur.nl/themas/digitale-duurzaamheid/>.
- Digital Preservation Coalition. Digital Preservation Handbook. <http://www.dpconline.org/handbook/>.
- Eszter Hargittai, Christian Sandvig. Digital Research Confidential: The Secrets of Studying Behavior Online. MIT Press, 2015.
- Annemieke de Jong. Digitale Preserving Beeld en Geluid: Beleid, Standaarden en Procedures. 2016. <http://publications.beeldengeluid.nl/pub/387>.
- Matthew G Kirschenbaum, Richard Ovenden, Gabriela Redwine, Rachel Donahue. Digital Forensics and Born-Digital Content in Cultural Heritage Collections. Citeseer, 2010. <https://www.clir.org/pubs/reports/reports/pub149/pub149.pdf>.
- Annette Markham, Elizabeth Buchanan. Ethical decision-making and internet research: Recommendations from the AoIR ethics working committee (version 2.0). 2012. <https://aoir.org/reports/ethics2.pdf>
- Jae Yun Moon, Lee Sproull. Essence of distributed work: The case of the linux kernel (originally published in volume 5, number 11, november 2000). First Monday, 2005.
- NISO Press. Understanding metadata. National Information Standards, 20, 2004. <http://www.niso.org/publications/press/UnderstandingMetadata.pdf>.
- Andrew S. Tanenbaum. Structured computer organization. Prentice Hall PTR, 1984.
- UNESCO. Charter on the Preservation of Digital Heritage. http://portal.unesco.org/en/ev.php-URL_ID=17721&URL_DO=DO_TOPIC&URL_SECTION=201.html.
- UNESCO/PERSIST Content Task Force. Richtlijnen voor de selectie van digitaal erfgoed voor langetermijnconservering. 2016. https://www.unesco.nl/sites/default/files/uploads/Comm_Info/finalpersistcontentguidelinesfinal_dutch.pdf
- Gaby Wijers, Hannah Bosma, onderzoek van de Culturele Coalitie Digitale Duurzaamheid naar born digital erfgoed. Generieke workflows born digital erfgoed. Behoud van born digital erfgoed in Nederland: film, fotografie, architectuur, kunst. http://ncdd.nl/site/wp-content/uploads/2015/10/CCDD-BornDigitalOnderzoek-def_2015.pdf. 2015.

Context DDS:

- Gerard Alberts, Ruth Oldenziel (Editors). Hacking Europe. From Computer Cultures to Demoscenes, Springer London, 2014.
- Marianne van den Boomen. Leven op het net: de sociale betekenis van virtuele gemeenschappen. Instituut voor publiek en politiek, 2000. <https://boom.home.xs4all.nl/boek/5kroegen.htm>.
- Geert Lovink. Dynamics of Critical Internet Culture. An Archive of Content Production (1994-2001). Institute of Network Cultures, Amsterdam 2009. http://networkcultures.org/_uploads/tod/TOD1_dynamicsofcriticalinternetculture.pdf.
- Howard Rheingold. The virtual community: Homesteading on the electronic frontier. MIT press, 2000. <http://www.rheingold.com/vc/book/>.
- Reinder Rustema. The Rise and Fall of DDS. Evaluating the ambitions of Amsterdam's Digital City. Amsterdam, 2001. <http://reinder.rustema.nl/dds/>.
- Menno van der Steen, Miro Schaap. Chat op het net. Rheydboek, Woerden, 1998.

Juridische en ethische aspecten:

- Archiefwet 1995. <http://wetten.overheid.nl/BWBR0007376/2015-07-18>.
- Autoriteit Persoonsgegevens. <https://autoriteitpersoonsgegevens.nl/nl/over-privacy/wetten/wet-bescherming-persoonsgegevens>.
- F.C.J. Ketelaar et al. Elke handeling telt. Archiefdiensten en de wet bescherming persoonsgegevens. <http://cf.hum.uva.nl/bai/home/eketelaar/WBP.pdf>. Dit is de in 2001 geactualiseerde versie van een artikel, verschenen in het Archiefblad 104//3 (mei 2000) 18-23 en 104/4 (juni 2000) 26-29.
- Nederlandse grondwet. Artikel 10: Privacy. <http://www.denederlandsegrondwet.nl/9353000/1/j9vvihlf299q0sr/vgrnbl6ah4zz>.
- Latanya Sweeney. k-anonymity: A model for protecting privacy. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 10(05):557{570, 2002. <http://www.cs.pomona.edu/~sara/classes/cs190-fall12/k-anonymity.pdf>.
- Wet bescherming persoonsgegevens. <http://wetten.overheid.nl/BWBR0011468/2017-03-10>.

Publicaties over het project:

- Gerard Alberts, Marc Went en Robert Jansma. Archaeology of the Amsterdam Digital City; why digital data are dynamic and should be treated accordingly". Internet Histories. Digital Technology, Culture and Society, pagina 1-14. April 2017. <http://www.tandfonline.com/doi/full/10.1080/24701475.2017.1309852>
- Tjarda de Haan. Project The Digital City Revives. A Case Study of Web Archaeology. Proceedings of iPRES2016, 13th International Conference on Digital Preservation, pagina 41- 45. Juli 2016. http://www.ipres2016.ch/frontend/organizers/media/iPRES2016/_PDF/IPR16.Proceedings_3_Web_Broschuere_Link.pdf.
- Tjarda de Haan, Paul Vogel. The reconstruction of The Digital City, a case study of web archaeology. Acoustic Space 14: Data Drift. Archiving Media and Data Art in the 21st Century, pagina 29-39. Oktober 2015. <http://www.bitsandbytesunited.com/?portfolio=publication-the-reconstruction-of-the-digital-city-a-case-study-of-web-archaeology>.



Verwijzingen naar project 'DDS Herleeft' van de partnerorganisaties:

Amsterdam Museum:

- <http://hart.amsterdammuseum.nl/re-dds>.
- Leesvoer voor de webarcheoloog: <https://hart.amsterdam/ddsleesvoer>.
- Project re:DDS in de media: <https://hart.amsterdam/nl/page/33445>.
- Video verslagen van het project DDS Herleeft: <https://www.youtube.com/user/reDeDigitaleStad>.
- Verzameling webbronnen voor ons project: https://del.icio.us/re_dds.

Nederlands Instituut voor Beeld en Geluid:

- <http://www.beeldengeluid.nl/kennis/projecten/de-digitale-stad-herleeft>.

Universiteit van Amsterdam:

- <http://www.uva.nl/shared-content/faculteiten/nl/faculteit-der-natuurwetenschappen-wiskunde-en-informatica/nieuws/2016/06/uva-en-vu-studenten-wekken-de-digitale-stad-tot-leven.html>.

Waag Society:

- <http://waag.org/nl/project/de-digitale-stad-herleeft>.

11

Bijlage II: Woordenlijst

Abstractieniveau:

Een abstractieniveau is een manier om over een digitaal object na te denken. Deze manier van denken wordt gebruikt om makkelijker over ingewikkelde objecten te kunnen praten. Neem bijvoorbeeld een schep. Het woord schep is een abstractieniveau. De schep bestaat uit een handvat, stok en schop. Deze onderdelen zijn een abstractieniveau lager van het object. We dalen nog een abstractieniveau als we naar de materialen kijken. Het is makkelijker om over een schep (als geheel) te praten, omdat iedereen meteen snapt wat je bedoelt. Ditzelfde doen we ook met computers. We praten bijvoorbeeld over programma's alsof het aparte fysieke objecten zijn, die los staan van de computer zelf, maar dit is niet waar. Het programma, dat bestaat uit nullen en enen, werkt omdat het zo op hardware staat. Op papier zou het niet werken. Wanneer we het over een schep hebben, kan de stok van hout of van metaal zijn. In beide gevallen hebben wij het, op het hogere abstractieniveau, over een schep. Er kunnen dus verschillen zijn (hout of metaal) die je op een hoger abstractieniveau niet ziet tussen de twee objecten. Zo lijkt een website voor een gebruiker hetzelfde op een Windows-computer of een Mac. Maar de onderliggende niveaus zijn niet hetzelfde. Het is een andere schep.

Checksums:

Checksums zijn soort vingerafdrukken die van bestanden gemaakt kunnen worden. Door de checksums te vergelijken kan gecontroleerd worden of twee bestanden hetzelfde zijn.

Digitaal erfgoed:

1. Gedigitaliseerd erfgoed: erfgoed dat van origine niet digitaal is, maar waarvan met digitalisering een reproductie is gemaakt.
2. Digitale informatie over erfgoed: bijvoorbeeld objectbeschrijvingen en contextuele informatie over het erfgoedobject. Deze gegevens worden opgenomen in collectieinformatiesystemen om de objecten goed te kunnen beheren of toegankelijk te maken.
3. Born digital erfgoed: erfgoedmateriaal dat van origine digitaal is, zoals foto's die met een digitale camera zijn gemaakt, digitale archieven, digitale kunst, websites, apps of games.

Emulatie:

Emulatie is het door middel van software nadoen van hardware en/of software. Virtuele Machines (VMs) zijn programma's die dit kunnen. Als je je (historisch) digitaal object weer werkend hebt gemaakt, kun je deze op een VM zetten. Door de VM-software blijven je object en alle onderliggende data toegankelijk.

Image:

Een image is een bit-identieke kopie van een opslagmedium. Het maken van een image voorkomt dat je (meta)data verliest. In een image zet je alle data over van de gehele datadrager, bit voor bit, inclusief de partities, boot manager en lege ruimtes. Op deze manier voorkom je problemen met compatibiliteit van filesystemen. Een image is dus een archiefbestand met álle data van het gekopieerde opslagmedium.

Metadata:

Dit is informatie over de data zelf. Gestructureerde gegevens over fysieke of digitale objecten. De drie belangrijkste soorten metadata zijn: beschrijvende metadata (nodig voor het identificeren en vinden van objecten), structurele metadata (deze leggen de relatie vast tussen individuele objecten die gezamenlijk een eenheid vormen), technische metadata (informatie over de vervaardiging van het object) en administratieve metadata (deze richten zich op beheer en management van objecten).¹⁰² Technische metadata zijn data die over een bestand gaan maar niet het bestand zelf zijn. Hieronder valt bijvoorbeeld de bestandsnaam, de datum wanneer het bestand is aangemaakt of gewijzigd, en alle andere vormen van gegevens die over een bestand gaan.¹⁰³

Metadata dating:

Door middel van de metadata kan achterhaald worden wanneer een bestand is aangemaakt en wanneer het voor het laatst gewijzigd is. Hiermee kun je dus achterhalen hoe oud een bestand is: Een soort carbon dating voor born digital erfgoed.

Nieuwsgroepen:

Nieuwsgroepen (usenet) zijn communicatiekanalen die al beschikbaar waren voor het World Wide Web. Qua structuur komen ze erg overeen met moderne fora: er zijn onderwerpen met daarin soms subonderwerpen waar vervolgens berichten in zitten. Gebruikers kunnen posts lezen en erop reageren. Een fundamenteel verschil voor nieuwsgroepen is dat het gedecentraliseerde systemen zijn. Ze worden niet gehost op een centrale server maar worden juist via peer-to-peer systemen verspreid.

Tools (software):

In de context van dit handboek zien wij tools als ieder programma of commando dat gebruikt kan worden om met historische data om te gaan. Voorbeelden hiervan zijn oudere versies van webbrowsers zoals Netscape, maar ook moderne besturingssystemen waarmee je de data benadert zoals grep, een Unix command line-programma dat kan zoeken door bestanden en mappen.

¹⁰² <http://www.den.nl/abc/Metadata/> en <http://datasupport.researchdata.nl/start-de-cursus/iii-onderzoeksfase/datadocumentatie/>

¹⁰³ National Institute Standards Organisation (NISO) (2004). Understanding Metadata. NISO press, USA (Online). <http://www.niso.org/publications/press/UnderstandingMetadata.pdf> (November 2011).

Webarcheologie:

Webarcheologie is een nieuw gebied in e-cultuur waarbij we relatief nieuwe (born digital) objecten die nog niet zo lang geleden verloren lijken te zijn gegaan, opgraven en reconstrueren met nieuwe (digitale) gereedschappen. Zowel de materie als de methoden om ons digitale verleden op te graven en te reconstrueren zijn piepjong en nog volop in ontwikkeling. Data vormen de nieuwe klei, scripts de nieuwe schoppen en het web is de allerjongste kleilaag die we gaan ontginnen. Webarcheologie bestaat uit een aantal processen, plannen, graven, reconstrueren en ontsluiten, die we in dit 'DIY (Do It Yourself) Handboek Webarcheologie' beschrijven.

TIP: Goede websites met verklarende woordenlijsten en/of leeromgevingen:

- 'ABC Digitaal Erfgoed' van Digitaal Erfgoed Nederland (DEN), nationale kennisinstituut voor digitalisering in de cultuursector. <http://www.den.nl/abc>.
- 'Digital Preservation Handbook' van de Digital Preservation Coalition <http://www.dpconline.org/handbook/glossary>.
- De cursus 'Leren Preserveren' van het Netwerk Digitaal Erfgoed, de NCDD en Het Nieuwe Instituut: <https://lerenpreserveren.nl/> en <https://lerenpreserveren.nl/woordenlijst/>.
- De website van The International Internet Preservation Consortium: <http://netpreserve.org/web-archiving/>.
- Platform voor domeinoverstijgende kennisuitwisseling en bewustwording rondom digitale duurzame toegang. Nationale Coalitie Digitale Duurzaamheid (NCDD) maakt deel uit van het Netwerk Digitaal Erfgoed. <http://www.ncdd.nl/>.
- Catalogus Duurzaamheidsbeleid. De Catalogus is opgezet en ingericht door de Nationale Coalitie Digitale Duurzaamheid (NCDD) in samenwerking met het Netwerk Digitaal Erfgoed: <http://wiki.ncdd.nl>.

12

Bijlage III:

Webarcheologische tools

Een overzicht van webarcheologische tools, per fase:

Plannen (vooronderzoek):

- Content management systeem voor je communicatieplatform - Een webapplicatie om kennis, documenten en gegevens te kunnen publiceren en delen.
- Full page screen capture, deze vind je online; ook als plugin in browsers - Tool die screenshots maakt van volledige webpagina's voor het archiveren van relevante bronnen voor het geval dat ze offline gaan.
- Wayback Machine, <https://archive.org/web/> - De Wayback Machine heeft snapshots van websites door de tijd heen. Deze kunnen gebruikt worden om het object te bestuderen.
- Webrecorder, <https://webrecorder.io/> - Een open source archiving tool die, met beperkingen, ook snapshots van dynamische content kan maken.
- Whois, <https://www.sidn.nl/whois> - Whois kan gebruikt worden om op te zoeken wie een domein bezit.
- Whois History, zie de 'Whois History' op websites zoals <http://whois.domaintools.com> - Whois history laat zien wie in het verleden de eigenaar van een domein was.

Graven (overzetten data):

- Checksums - Checksums kunnen gebruikt worden om relatief snel te controleren of het overzetten van data goed gelukt is.
- Images, zie <https://www.youtube.com/watch?v=jrJTQF3o5c4> - Een image is een bit-identieke kopie van een opslagmedium. Hiermee voorkom je verlies van (meta)data.

Reconstructie (data verkennen):

- Bitcurator, <https://www.bitcurator.net/> - Een open source Linux-werkomgeving met verschillende tools die kunnen helpen bij het verkennen van datasets.
- De-NISTing, <https://github.com/tifkap/deNISTng> en <https://www.nist.gov/> - Tool voor het automatisch herkennen van bekende software in een filesysteem.
- Find - Find is een Unix command line tool voor het zoeken in folders.
- Grep - Grep is een Unix command line tool voor het zoeken in bestanden.
- Metadata dating - Metadata dating is een methode waarbij de metadata van bestanden wordt gebruikt om te bepalen wanneer bestanden zijn aangemaakt.
- Op maat gemaakte tools van de History of Digital Cultures projecten aan de Universiteit van Amsterdam - Verschillende custom tools gemaakt door de studenten die de DDS data te verkennen en ontsluiten. Voorbeelden hiervan zijn data visualisatie tools en op maat gemaakte SQL-databases.

Reconstructie:

- Op maat gemaakte tools van de History of Digital Cultures projecten aan de Universiteit van Amsterdam. Afhankelijk van doel en object, verschillende custom tools gemaakt door de studenten die de DDS data ontsluiten. Bijvoorbeeld De Replica van de DDS-website en een visualisatie van de e-mail data.
- Virtuele machines - Virtuele machines geven een hardware-onafhankelijke werkomgeving die door een heel team gebruikt kan worden.
- Wayback Machine, <https://archive.org/web/> - De Wayback Machine kan gebruikt worden om de reconstructies te vergelijken met hoe het toen was.

Ontsluiten:

- Checksums - Checksums zijn hashes, soort vingerafdrukken die van bestanden gemaakt kunnen worden. Door de hashes te vergelijken kan gecontroleerd worden dat twee bestanden hetzelfde zijn.
- Virtuele machines - Virtuele machines zijn makkelijk overdraagbaar naar een archief of museum en makkelijk te controleren op goede overdracht (met checksums) en platform onafhankelijk.
- Web ARChive (WARC) - Een veelgebruikt archiveringsformaat voor het opslaan van webcrawls.
- Webrecorder, <https://webrecorder.io/> - Een open source archiveringstool die naast de statische content ook bepaalde dynamische content kan archiveren.

13

Colofon

Auteurs:

Tjarda de Haan (gast-conservator digitaal erfgoed, webregisseur)
Robert Jansma (student en assistent-docent)
Paul Vogel (systeembeheerder, programmeur).

Eindredacteur:

Tjarda de Haan.

Uitgever en vormgeving:

Amsterdam Museum.

Vormgeving:

Studio September, Edo Mulder (Amsterdam Museum).

Met dank aan de teamleden van het Amsterdam Museum, Waag Society, Universiteit van Amsterdam en het Nederlands Instituut voor Beeld en Geluid, Carolien van Zuilekom van het Netwerk Digitaal Erfgoed (NDE) en de Nationale Coalitie Digitale Duurzaamheid (NCDD) en Kees Teszelszky van de Koninklijke Bibliotheek voor het nauwkeurig en kritisch meelesen.



Naamsvermelding-NietCommercieel-GelijkDelen

<https://creativecommons.org/licenses/by-nc-sa/4.0/deed.nl>

Amsterdam, 2017.



meer informatie: hart.amsterdam/freeze