



Surf naar mij! Stem voor mij!

Archiveren van websites bij de gemeenteraadsverkiezingen van 2012.



Filip Boudrez, Stadsarchief Antwerpen / eDAVID

Een Hollander die in de Schelde wordt geduwd. Havenschepen en lijstduwer Marc Van Peel die zijn duwkracht in de verf zet door al telefonierend de deur van het stadhuis gesloten te houden. Een muzikaal zangersduo pakt uit met een campagnelied en wil van het Antwerpse burgemeesterschap een duobaan maken. Het lanceren van een heuse verkiezingsapp waarmee de kiezer in de huid van de politieke kandidaat kruipt. Het zijn maar enkele voorbeelden van hoe partijen en kandidaten bij de recentste verkiezingen in hun campagnes het web en nieuwe technologieën gebruikten. Het stadsarchief Antwerpen archiveerde ze en nam ze op in het digitale depot waar ze duurzaam bewaard blijven voor de toekomst.

DIGITALE CAMPAGNES

In de aanloop van de gemeenteraadsverkiezingen van oktober 2012 viel op dat politieke partijen en hun kandidaten veel meer dan in 2006 het web en nieuwe technologieën gebruikten om campagne te voeren. Bij de vorige gemeenteraadsverkiezingen voerden ze hoofdzakelijk nog een papieren verkiezingscampagne en bleef het gebruik van websites voornamelijk beperkt tot initiatieven van buiten de politiek. Een Antwerps voorbeeld van dit laatste was 'Droom de stad', een initiatief van theatermakers, gezelschappen en

andere organisaties waarbij dromen en wensen voor de stad online werden ingezameld en gepubliceerd. De evolutie naar meer digitale campagnevoering tekende zich al af in aanloop van de volksraadpleging over de Oosterweelverbinding in 2009. Actiegroepen, scholen, burgercomités, enz. maakten toen al intensief gebruik van websites, filmpjes en animaties om pro of contra het ontwerp voor de Oosterweelverbinding campagne te voeren.

In Antwerpen was de evolutie naar meer digitale campagnes in de verkiezingsaanloop heel

goed merkbaar. Meerdere factoren bieden hier een verklaring voor. Algemeen geldt dat de opmars van het web zich almaar verder zet en dat internettoepassingen niet worden meegeteld als verkiezingsuitgave. Maar in Antwerpen speelden ook nog andere factoren een rol. Aanvankelijk was er een verbod op politieke aanplakborden zodat partijen en kandidaten wel (digitale) alternatieven moesten zoeken. En natuurlijk speelde ook de titanenstrijd tussen Patrick Janssens en Bart De Wever een rol. Wie aandacht wou, moest uitpakken met opvallende initiatieven. De persoonlijke

magazines¹ deden hun intrede, maar bovenal werd gebruik gemaakt van websites en digitaal campagnemateriaal.

VAN CAMPAGNEFILMPJES TOT EEN VERKIEZINGSAPP

Elke partij maakte gebruik van websites om campagne te voeren. De wijze waarop ze dit deden, verschilde van partij tot partij. Het brede scala van beschikbare webtechnologieën werd hierbij ingezet. De grotere partijen lieten een volledig nieuwe website bouwen of creëerden een subsite voor de website van de nationale partij. Open VLD presenteerde

zijn standpunten niet zomaar in de vorm van een klassieke website, maar maakte een site op basis van Google Maps zodat pinpoints met programmapunten meteen aangaven over welke geografische locatie in Groot-Antwerpen het ging. De kleinere partijen gebruikten blogs of kant-en-klare CMS-platformen. Lijsttrekkers en lijstduwers hadden meestal ook een persoonlijke website. En natuurlijk wierpen de meeste kandidaten ook hun Facebookpagina mee in de kiesstrijd of maakten ze een fanpagina op Facebook.

Het bleef echter niet bij het maken van websites of webpagina's. Een nieuw verschijnsel was het maken van campagne-filmpjes. Havenscheper Marc Van Peel haalde de nationale media met zijn ludieke filmpjes die op YouTube en Vimeo werden gepost en lokte hiermee



Patrick Janssens en Meyrem Almaci brachten een eenmalig tijdschrift uit tijdens de campagne voor de gemeenteraadsverkiezingen 2012 in Antwerpen.

veel bezoekers naar zijn website². Partijgenoot Philip Heylen aanhoorde dezelfde arrogante Hollander op de panoramaverdieping van het MAS³. Gerolf Annemans haalde voor zijn filmpje niet alleen inspiratie bij Marc Van Peel, maar ook bij Bob Dylan⁴. De VRT organiseerde zelfs een actie waarbij lokale kandidaten hun filmpje konden opladen op de VRT-website. Voornamelijk districtskandidaten maakten hier gebruik van.

De muzikale broers Mathieu en Guillaume bouwden hun website en hun campagne op rond een liedje. Sp.a-boegbeeld Robert Voorhamme ging nog een stap verder en liet een heuse verkiezingsapp ontwikkelen: de 'Doe de Voorhammenizer'. Met de app kunnen gebruikers hun foto opladen, de bril van Voorhamme opzetten en een gepaste slogan kiezen. Deze app werd beschikbaar gesteld in Flash op zijn website, maar kon ook als app voor iOS en Android worden gedownload via iTunes en Google Play.

Het gebruik van sociale media bleef hoofdzakelijk beperkt tot Facebook. Speciale twitteraccounts voor de verkiezingen vonden we niet terug. Wel werd vanop bestaande en persoonlijke twitteraccounts volop getweet over de verkiezingen en de campagnes.

ARCHIVERINGSBELEID

Het archiveren van politiek campagnemateriaal voor de verkiezingen maakt sinds jaar en dag deel uit van het acquisitieprofiel van het stadsarchief Antwerpen. Dit campagnemateriaal heeft nu steeds meer een digitale vorm.

Het stadsarchief heeft daarnaast ook een jarenlange traditie in het archiveren van websites. Het onderzoek naar geschikte archiveringsmethoden en -strategieën werd gestart binnen het DAVID-project en werd vanaf 2002 in de praktijk gebracht. Zo werd de eerste website

van de stad uit 1995 succesvol gearcheveerd⁵ zodat die nog steeds raadpleegbaar is via de website van het FelixArchief⁶. Het stadsarchief Antwerpen archiveert niet alleen de vele websites van de stad Antwerpen⁷, maar archiveert ook de websites van evenementen en speciale campagnes. Concrete voorbeelden in de context van verkiezingen zijn de website van het initiatief 'Droom de stad' (2006)⁸ en de vele websites van de actiegroepen naar aanleiding van de volksraadpleging van 2009. Toen werden de websites van actiegroepen zoals Ademloos⁹, Straten-Generaal¹⁰, Rapper met de Wapper¹¹, enz. gearcheveerd en in het digitale depot opgenomen.

Voor het digitale campagnemateriaal houdt dit onder meer in dat in aanloop van de verkiezingen actief naar websites en filmpjes wordt gezocht. Er werd gezocht naar de websites van de lokale afdelingen van politieke partijen en van lokale lijsten en hun kandidaten. De speurtocht op het web werd aangevuld met het intensief volgen van de nationale en lokale media, waarin volop werd bericht over de campagnevoering.

Na het in kaart brengen van digitaal campagnemateriaal startte het contacteren van de rechthebbenden zodat de vereiste toestemming voor archivering en vooral online beschikbaarstelling via de website van het stadsarchief Antwerpen kon worden verkregen. Op het tijdstip van archivering had het stadsarchief Antwerpen diverse archiveringsacties bij de kabinetten van burgemeester en schepenen lopen zodat van deze contacten gebruik kon worden gemaakt.

MOMENTOPNAMES MAKEN

Voor het archiveren van de websites en het aanverwante digitale campagnemateriaal past het stadsarchief

“VOOR HET ARCHIVEREN VAN DE WEBSITES EN HET AANVERWANTE DIGITALE CAMPAGNE-MATERIAAL PAST HET STADSARCHIEF ANTWERPEN DE METHODEN EN ‘BEST PRACTICES’ TOE DIE BINNEN eDAVID WERDEN ONTWIKKELD.”



Een voorbeeld van een video op de website van Rood!Antwerpen die inmiddels op Youtube niet meer beschikbaar is. De video werd gelukkig afzonderlijk vastgelegd en in het digitale depot van het FelixArchief gearchieveerd.

Antwerpen de methoden en best practices toe die binnen eDAVID werden ontwikkeld¹². Al naargelang het type campagnemateriaal wordt een andere methode gebruikt.

WEBSITES

Het maken van een momentopname van een gewone website gebeurt door middel van een webharvester of een crawler. Dat is een computerprogramma waarbij een volledige website of een geselecteerd deel van een website op schijf wordt vastgelegd. Het resultaat is een momentopname op een specifiek tijdstip. Deze methode is vergelijkbaar met het maken van een foto van een bepaalde situatie. Er zijn veel tools voor het vastleggen van een website beschikbaar¹³. Het stadsarchief Antwerpen gebruikt doorgaans de tools HTTrack¹⁴, de Web Curator Tool¹⁵ of Wget om een website vast te leggen. Al naargelang de website en het resultaat gebruiken we een andere tool. Op basis van de technische kenmerken van een website kunnen we doorgaans wel inschatten welke tool het beste resultaat zal opleveren, maar soms is de keuze van de gepaste tool en de gebruikte instellingen ook een kwestie van trial-and-error.

Het resultaat van het zo'n vastleggingsoperatie in de

vorm van een momentopname is maar zelden meteen perfect of een exacte archiveringskopie van de online website. De momentopnames blijken dikwijls een aantal fouten te bevatten of er blijken een aantal componenten te ontbreken. Een goede kwaliteitscontrole waarbij de online website en de geregistreerde website met elkaar worden vergeleken, is dus belangrijk. Naast de visuele controle worden een aantal zaken systematisch en geautomatiseerd gecontroleerd: de aanwezigheid van alle afbeeldingen, de geldigheid van alle interne links, enz. Het stadsarchief Antwerpen gebruikt voor deze laatste controles een eigen ontwikkelde tool waarmee eventuele mankementen snel kunnen worden opgespoord en (geautomatiseerd) gecorrigeerd. Soms is ook nog een manuele correctie vereist. Enige kennis van HTML en CSS is hierbij vereist. Is het aantal fouten te hoog, dan wordt met een andere tool of met andere instellingen een nieuwe momentopname gemaakt.

Het resultaat van zo'n momentopname is een grote set van gelinkte digitale objecten. Het bestandsformaat van deze digitale objecten kan heel verscheiden zijn. Naast de te verwachten formaten van webtechnologieën zoals (X)HTML, CSS, JavaScript, Flash, enz. komen in zo'n momentopname nog heel andere formaten voor. Op de meeste websites wordt immers nog informatie in allerhande formaten aangeboden: PDF, Word, Excel, Powerpoint, audio, video, enz. Deze digitale inhoud vormt een bijzondere uitdaging op het vlak van duurzaamheid en vastlegging als onderdeel van de momentopname.

YOUTUBE EN VIMEO

Dit laatste stelt zich in het bijzonder ten aanzien van digitale video op websites. Bij de meeste websites wordt digitale video door middel van kanalen als YouTube en Vimeo

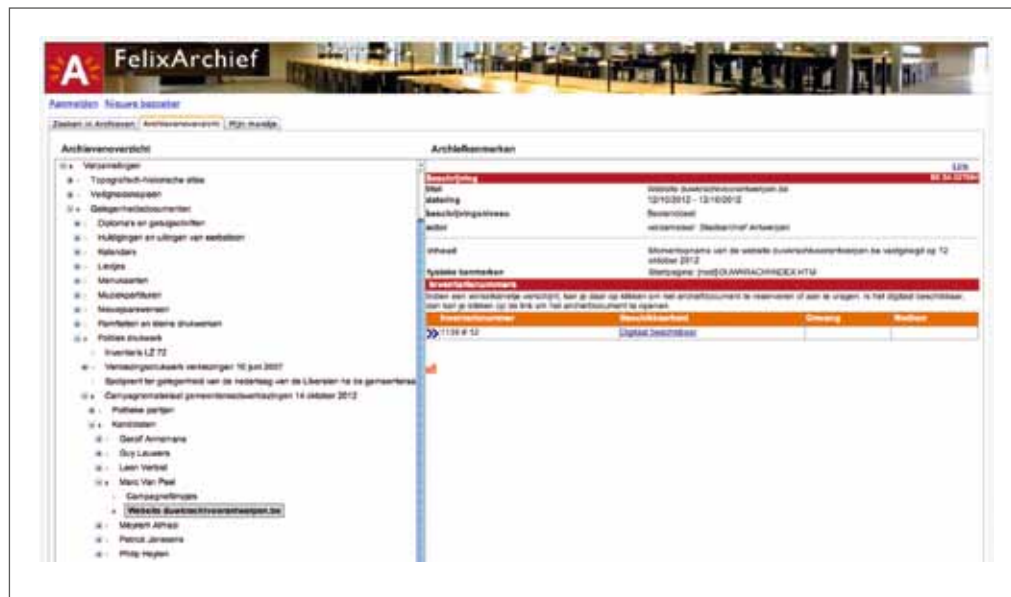
beschikbaar gesteld. Deze video's ontbreken doorgaans in de momentopnames die door middel van een gewone webharvester worden vastgelegd. In de momentopnames blijven externe links naar de online video's staan. En deze links blijven maar operationeel zolang de video's online beschikbaar blijven op YouTube en Vimeo. Hierdoor ontstaat het risico dat op een gegeven moment een inhoudsonderdeel van de website ontbreekt. De digitale video's worden bijgevolg best ook afzonderlijk vastgelegd.

Voor het vastleggen van online video's op kanalen als YouTube en Vimeo zijn diverse tools beschikbaar. De archivaris kan gebruik maken van een webbased online tool of kan hiervoor een computerprogramma installeren. Een bekende webbased tool voor het vastleggen van YouTube-video's is Keepvid¹⁶. Voor Vimeo zijn ook diverse online services beschikbaar¹⁷, maar kan ook afzonderlijke software worden geïnstalleerd¹⁸. Wij kiezen ervoor om de digitale video meteen in een geschikt archiveringsformaat vast te leggen (bijv. MPEG-4) en vermijden het gebruik van Flash video.

Naast het vastleggen van de digitale video als afzonderlijk videobestand is het minstens even belangrijk dat de locatie van de video binnen de website en de link naar het externe videobestand goed worden gedocumenteerd. Dit is zeker belangrijk wanneer meerdere externe video's op de website staan. Mocht op een gegeven tijdstip de online video niet langer beschikbaar zijn, dan moet de webpagina met behulp van de extra documentatie opnieuw reconstrueerbaar zijn.

GOOGLE MAPS

De problematiek van het archiveren van Google Maps toepassingen binnen websites is vergelijkbaar met die van het archiveren van externe video's: in de momentopname



Gearchiveerde websites en ander digitaal campagnemateriaal worden beschreven in het archievenoverzicht van het stadsarchief Antwerpen. Technische bijzonderheden en eventuele correcties worden gedocumenteerd in de ISAD(G)-beschrijving. De technische kenmerken van elk digitaal object worden gedocumenteerd in de index AIP's die mee in het digitale depot worden bewaard.

van de website ontbreekt de Google map en zijn pinpoints. De vastgelegde webpagina bevat enkel de link naar de online Google map.

Er zijn verschillende methoden om de Google map en zijn pinpoints op een afzonderlijke wijze vast te leggen. De kaarten zelf zijn in essentie niets meer dan afbeeldingen in GIF-, PNG- of JPEG-formaat. Deze kunnen als afzonderlijke afbeeldingen worden vastgelegd. Ook voor de archivering van de pinpoint-informatie zijn meerdere opties. Een eerste mogelijkheid is het bewaren van de Google map in de vorm van een KML-bestand. Enige kennis van de Google map API is hiervoor vereist. In de webpagina vind je de nodige parameters voor de kaart terug, maar de URL voor het oproepen van de kaart en het opslaan als KML-bestand dien je wel manueel samen te stellen. Een tweede mogelijkheid is het archiveren van het XML-bestand met de gegevens van de Google map en zijn pinpoints. In de broncode van de webpagina vind je de naam van het XML-bestand op de webserver terug die deze gegevens bevat. Dit XML-bestand bevat de markers, de locatie, de titel en de beschrijving van de pinpoints. Dit XML-bestand kun je gewoon

oproepen in een webbrowser en als een afzonderlijk XML-document bewaren.

FACEBOOK

Hoewel Facebookpagina's of fanpagina's in essentie webpagina's zijn, levert het maken van een momentopname met de beschikbare webharvesters geen goede resultaten op. Het resultaat blijkt doorgaans onvolledig te zijn en te veel informatie van andere Facebookaccounts te bevatten. Eén van de redenen hiervoor is dat Facebook regelmatig wijzigingen in zijn software aanbrengt zodat het archiveren van Facebookpagina's een kat-en-muisspelletje is tussen Facebook en ontwikkelaars van archiveringssoftware.

Om een volledige momentopname van een Facebookaccount te maken, kan beter de downloadfunctie van Facebook zelf worden gebruikt¹⁹. Via de beheerderspagina's van het Facebookaccount kan een 'kopie van de Facebookaccount' worden aangevraagd. De beheerder ontvangt vervolgens op zijn e-mailadres een bericht met downloadlink naar een zip-bestand die alle informatie (incl. afbeeldingen) van zijn Facebookaccount bevat. Deze methode is echter niet perfect. Het voornaamste

nadeel van deze 'content dump' is het ontbreken van de oorspronkelijke 'look-and-feel' van de webpagina's. Een ander nadeel is dat deze functie niet beschikbaar is voor fanpagina's op Facebook en dat profielfoto's van vrienden ontbreken.

Om wel de oorspronkelijke 'look and feel' van de webpagina's te behouden en om fanpagina's te archiveren kan beter de ArchiveFB-addon voor de Firefox webbrowser worden gebruikt²⁰. Dit is een addon waarmee de publieke webpagina's van een Facebookaccount of -fanpagina in hun oorspronkelijke look-and-feel wordt vastgelegd. Voor een goede archivering worden beide methodes idealiter met elkaar gecombineerd.

ARCHIVEREN BESCHRIJVEN

De gearchiveerde websites en het aanverwante digitale campagnemateriaal krijgen een ISAD(G)-archiefbeschrijving die wordt toegevoegd aan het archievenoverzicht van het stadsarchief Antwerpen (<http://zoeken.felixarchief.be>). Bij het beschrijven van gearchiveerde websites wordt bijzondere aandacht besteed aan het documenteren van de datum van momentopname

**“HET ARCHIVEREN
VAN WEBSITES
MAAKT
INTEGRAAL DEEL
UIT VAN HET
ARCHIVERINGS-
BELEID VAN HET
STADSARCHIEF
ANTWERPEN.”**

(= datum van het archiefdocument), de URL, de technische kenmerken en de afhankelijkheden van de website. Zo zijn websites dikwijls specifiek voor een bepaalde webbrowser(versie) ontwikkeld of werken ze niet goed in bepaalde webbrowsers. Deze kenmerken worden expliciet vermeld in de archiefbeschrijving. Ook de startpagina van de website wordt expliciet geregistreerd in de archiefbeschrijving. Eventuele correcties en aanpassingen in de momentopname worden gedocumenteerd zodat gebruikers weten welke belangrijke bewerkingen werden uitgevoerd.

OPEN FORMAAT

De gearchiveerde websites worden in een zo open mogelijk formaat gearchiveerd. Vanwege deze reden maken we bewust geen gebruik van het WARC-formaat voor de archivering van websites. WARC is een containerformaat voor gearchiveerde websites waarbij de verschillende dataobjecten die samen een website vormen in één of meerdere containerbestanden worden verpakt²¹. We kiezen niet voor WARC vanwege twee redenen: een WARC-bestand bevat geen *representation information* over

de opgeslagen dataobjecten en deze metadata bijkomend registreren op basis van WARC-bestanden is nodeloos complex. Ten tweede willen we voor de duurzame archivering in het digitale depot zo weinig mogelijk afhankelijk zijn van extra of overbodige reconstructieschakels. In de plaats daarvan archiveren we websites zo open mogelijk als een set van gelinkte digitale objecten zodat ook voor de raadpleging geen andere software dan een gewone webbrowser nodig is. Grote websites kunnen hierdoor wel uit tienduizenden digitale bestanden bestaan, maar de inrichting van het digitale depot van het stadsarchief Antwerpen is hier op voorzien. Het digitale depot heeft geen beperking op aantal bestanden of omvang van de bestanden per inventarisnummer.

METADATA REGISTREREN

De onderlinge samenhang tussen de gelinkte digitale objecten wordt wel expliciet geregistreerd in de vorm van een index AIP. Net zoals voor alle andere digitale en gedigitaliseerde archieven wordt voor elke gearchiveerde website een index AIP gegenereerd. Deze AIP-klasse bevat niet alleen een index van de (sub-) mappen en hun bestanden,

maar bevat voor een gearchiveerde website ook alle technische metadata van en over de digitale objecten die nodig zijn ter ondersteuning van de langetermijnarchivering. Als onderdeel van deze technische metadata worden de checksums (*fixity information*) en de essentiële karakteriserende informatie (*representation information*) geregistreerd. Op die manier archiveren we niet zomaar dataobjecten, maar informatieobjecten en brengen we het OAIS-informatiemodel in de praktijk.

Het registreren van deze metadata verloopt volledig geautomatiseerd en is een onderdeel van het opnameproces in het digitale depot.

RAADPLEGINGSKOPIEËN

De laatste stap in het archiveringsproces is het genereren van een raadplegingskopie voor gearchiveerde websites. Voor de raadpleging van de gearchiveerde websites via de website van het stadsarchief Antwerpen stellen we een ZIP-bestand beschikbaar. Dit ZIP-bestand wordt na opname in het digitale depot automatisch gegenereerd door ons archiefbeheerssysteem MACZ.

BESLUIT

Het archiveren van websites maakt integraal deel uit van het archiveringsbeleid van het stadsarchief Antwerpen. Binnen de reguliere werkprocessen en de verwerving van nieuwe archieven is het archiveren van websites een inherent onderdeel. Ook bij evenementen zoals verkiezingen wordt bijzondere aandacht besteed aan de archivering van websites. Dit laatste vormt een uitdaging want dit vraagt doorgaans snelle acties. Door de eDAVID-methodes voor websitesarchivering toe te passen, slaagt het stadsarchief Antwerpen er in om de talrijke websites en de daarbij horende media vast te leggen, te documenteren, te beschrijven en op te nemen in het digitale depot. ■■

1 Voorbeelden hiervan zijn o.m. de magazines van Patrick Janssens (*Patrick*), Meyrem Almaci, Mathieu en Guillaume en Vlaams Belang (*Zot van A*).

2 "Een Hollander en een Antwerpenaar staan samen aan de Schelde. Zegt de Hollander:" (http://www.youtube.com/watch?v=_RgI0JOTwdg) en "Even Van Peel bellen..." (http://www.youtube.com/watch?v=K_k0T9JBkOc)

3 "Een Hollander en een Antwerpenaar staan op het dak van het MAS. Zegt de Hollander:" (<http://www.youtube.com/watch?v=rmRRq34DNUY>)

4 <http://www.youtube.com/watch?v=yVqIGRYC510>

5 F. Boudrez, *Van backup tot gearchiveerde website. De archivering van de eerste versies van de Digitale Metropool Antwerpen*, Antwerpen, 2002. (<http://www.edavid.be/davidproject/teksten/DAVIDbijdragen/Recuperatie.pdf>)

6 http://zoeken.felixarchief.be/zHome/Home.aspx?id_isad=1209

7 http://zoeken.felixarchief.be/zHome/Home.aspx?id_isad=237

8 http://zoeken.felixarchief.be/zHome/Home.aspx?id_isad=9851

9 http://zoeken.felixarchief.be/zHome/Home.aspx?id_isad=197159

10 http://zoeken.felixarchief.be/zHome/Home.aspx?id_isad=197163

11 http://zoeken.felixarchief.be/zHome/Home.aspx?id_isad=197165

12 Op de website van het Expertisecentrum DAVID zijn meerdere rapporten, 'best practices' en aanbevelingen gepubliceerd. Zie <http://www.edavid.be/publicaties.php#3>

13 Een uitgebreid overzicht met beschrijving van de beschikbare tools is beschikbaar op <http://netpreserve.org/web-archiving/tools-and-software>.

14 <http://www.httrack.com>

15 <http://webcurator.sourceforge.net>

16 <http://www.keepvid.com>

17 Bijv. <http://downloadvimeo.com>, <http://www.videograbber.net/free-vimeo-downloader>, <https://chrome.google.com/webstore/detail/vimeo-download-videos>

18 Bijv. 4K Vimeo Downloader

19 Deze functie is beschikbaar voor beheerders van facebookaccounts op de pagina met accountinstellingen. Onder de algemene accountinstellingen is de functie 'Een kopie downloaden van je Facebook-gegevens' beschikbaar.

20 <https://addons.mozilla.org/en-US/firefox/addon/archivefacebook>. Zie ook M. Kelly, *Archive Facebook*, 2011.

21 Voor meer informatie over het WARC-formaat zie: F. Boudrez, *Technische standaarden voor digitale archiefdocumenten*, in: *Archiefbeheer in de praktijk*, 5714, p. 79-80.