

Een web van webarchieven

Een pleidooi voor samenwerking in Nederland

Marcel Ras en Arnoud Goos ■

Het internet is niet meer weg te denken uit ons dagelijks leven. Maar het web is vluchtig en de informatie waarop je vandaag je beslissing baseert, kan morgen verdwenen zijn. Webarchivering is daarom van groot belang, maar vindt nog te weinig en te versnipperd plaats. Tijd voor actie dus.

a

Voor een grote hoeveelheid online informatie geldt dat het wenselijk en soms zelfs noodzakelijk is dat deze toegankelijk blijft, ook nadat het offline wordt gehaald. De redenen hiervoor zijn divers: naast cultuurhistorische gronden (collecties van culturele instellingen, maar ook sociale media die veel inzicht geven over het tijdbeeld) gaat het ook om archiefwettelijke plichten (overheidsinformatie), wetenschappelijke redenen, bedrijfsbelangen en juridische en journalistieke overwegingen. Deugdelijke archivering van het web voor de lange termijn is dus cruciaal, maar op dit moment worden naar schatting maar 0,2% van alle sites in het Nederlandse domein bewaard. Hoewel veel instellingen in Nederland van het belang van webarchivering overtuigd zijn en er een aantal goede initiatieven lopen, is er landelijk gezien dus nog veel te winnen op dit gebied.

Webarchivering in Nederland

De eerste Nederlandse website kwam in februari 1992 online. In de 25 jaar die volgden, explodeerde het web tot de 5,7 miljoen Nederlandse domeinnamen die het .nl-domein nu telt en een onbekend aantal Nederlandse sites met een andere extensie (variërend van .com, .eu tot .info of zelfs .amsterdam). Daarmee is het .nl-domein een van de grootste ter wereld. Het archiveren van het Nederlandse web kwam echter pas relatief laat op gang. The Internet Archive begon al in 1996 met het archiveren van websites wereldwijd en de nationale bibliotheken van onder andere Groot-Brittannië, Australië, Denemarken en IJsland volgden al snel met het *harvesten* van websites binnen hun respectievelijke domeinen. Zij zagen al gauw de noodzaak van samenwerking en richtten daarvoor in 2003 de International Internet Preservation Consortium (IIPC) op. Het eerste Nederlandse webarchiveringsproject kwam pas in januari 2000 van de grond, bij het Documentatiecentrum Nederlandse Politieke Partijen (Archipol). Na een kort experiment in 2003 startte de Koninklijke Bibliotheek in 2006 met het archiveren van Nederlandse websites. Nu, ruim tien jaar later,

bevat die webcollectie een dwarsdoorsnede van het Nederlandse web met bijna 12.000 websites.¹

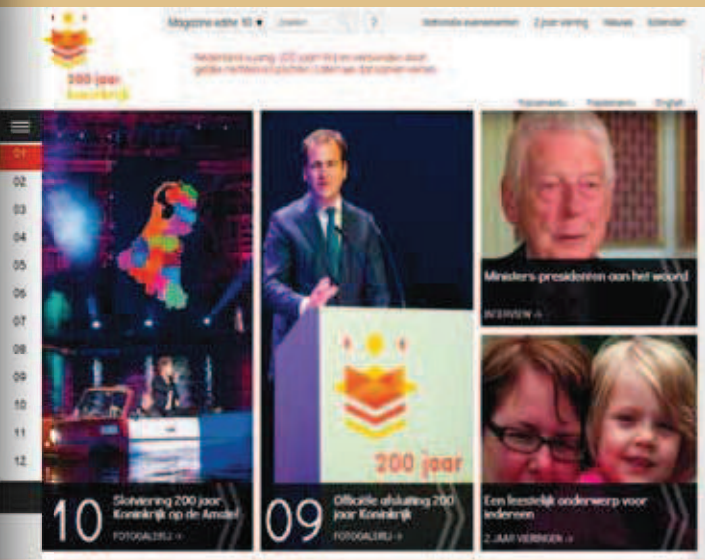
Na 2006 ontstonden er diverse initiatieven. Zo archiveert Beeld en Geluid mediagerelateerde websites (van bijvoorbeeld omroepen, producenten of fan-forums) en richt het Stadsarchief Rotterdam zich op websites over het leven in de stad Rotterdam. Het RHC Groninger Archieven bewaart websites uit de provincie Groningen en het Nationaal Archief is gestart met het archiveren van de websites van de rijksoverheid.

Meer dekking en meer coördinatie

Hoewel steeds meer gemeenten, overheidsinstellingen, musea en kennisinstellingen starten met het archiveren van de delen van het internet die voor hen relevant zijn, vormen zij nog steeds een relatief kleine groep. Ook daar waar webarchivering



Verdwenen websites. Van linksboven naar rechtsonder: de website van de omroep RKK (collectie Beeld en Geluid), de website van het Geldmuseum (collectie Koninklijke Bibliotheek – KB), de website van warenhuis V&D (collectie KB) en de website van Schoolbieb (collectie KB).



verplicht is gesteld door de Archiefwet loopt de praktijk nog behoorlijk achter. De Erfgoedinspectie constateerde eind 2016 dat de honderden websites waarmee de centrale overheid de burger tegemoet treedt nauwelijks worden gearchiveerd.² Slechts 2% van de betreffende organisaties archiveert de eigen websites. Het percentage organisatieonderdelen van de rijksoverheid dat de sociale media bewaart, is nog kleiner. De burger of het bedrijf dat zijn recht wil halen op basis van eerder verstrekte online informatie loopt dus een groot risico om achter het net te vissen.

Het daadwerkelijk doen van webarchivering in Nederland wordt belemmerd doordat organisaties niet over de benodigde kennis en kunde beschikken (hoe begin je, welke tools en technieken zijn er, kun je dit uitbesteden of moet je zelf aan de slag?). Daarnaast ontbreekt het in Nederland ook aan samenwerking: een gecoördineerde aanpak waarbij organisaties kennis en ervaringen uitwisselen en taken verdelen of afstemmen. Ondanks dat webarchieven vaak verschillende beweegredenen en een verschillende aanpak hebben, zijn de problemen waar ze tegenaan lopen veelal hetzelfde. Door kennis te delen en werkzaamheden op elkaar af te stemmen, wordt het makkelijker om (technische of juridische) struikelblokken aan te pakken.

Selectiebeleid

In sommige landen is er binnen een wettelijk kader een zogenaamde *domaincrawl* mogelijk, waarbij alle websites van het landelijke domein (in ons geval .nl) gearchiveerd worden. In Nederland is zo'n *domaincrawl* juridisch (maar ook technisch en financieel) onmogelijk, omdat er voorafgaand aan het archiveren van een website een bepaalde mate van toestemming nodig is van de rechthebbenden. Webarchivering is in ons land dus in alle gevallen gebaseerd op een specifiek selectiebeleid. Ieder initiatief bepaalt zelf welke websites er gearchiveerd worden en dat werkt samenwerking en coördinatie niet in de hand.

Webarchivering in Nederland zou dus enorm geholpen zijn met

enerzijds meer kennis en kunde bij de instellingen en anderzijds meer afstemming tussen die partijen. Recent zijn verschillende initiatieven opgestart die hieraan bijdragen.

Coördinatie webarchivering Nederland

Begin 2016 startte binnen het werkpakket Digitaal Erfgoed Houdbaar van het Netwerk Digitaal Erfgoed (NDE) het project 'Coördinatie van webarchivering in Nederland'. Dit project wordt geleid door Beeld en Geluid onder coördinatie van de Nationale Coalitie Digitale Duurzaamheid (NCDD). De eerste concrete activiteit was het inventariseren van het huidige aantal en het soort webarchieven in Nederland, met als doel toekomstige samenwerking te vergemakkelijken. Dit overzicht zal resulteren in een nationaal register voor webarchieven, waarvoor in dit project een opzet wordt gemaakt. De doelstelling is dat in dit register voor zowel professionals als het brede publiek op één plek inzichtelijk is welke websites op welke plek gearchiveerd zijn. Ook moet worden bijgehouden sinds wanneer en met welke interval deze websites worden gearchiveerd, met welke software dit wordt gedaan en met welk doel een website is gearchiveerd. Dit register kan worden doorontwikkeld tot een nationaal portaal dat toegang biedt tot alle gearchiveerde websites in het Nederlandse domein. In eerste instantie zal dit – door juridische belemmeringen – waarschijnlijk vooral gaan om toegang tot de metadata ervan, inclusief informatie over de locatie van de gearchiveerde websites.

Expertgroep webarchivering

Om de landelijke samenwerking verder richting te geven, heeft de NCDD eind 2016 een expertgroep webarchivering ingericht. Deze expertgroep bestaat uit vertegenwoordigers van de verschillende domeinen in het werkterrein van de NCDD (archieven, bibliotheken, musea, media/audiovisueel en wetenschap) en heeft als missie de praktische samenwerking en kennisdeling te bevorderen bij het archiveren van het Nederlandse web.

>>



Studiedag 'Een web van webarchieven' op 17 november 2016. Foto: NCDD.

- >> Het (her)gebruik en de toegang op de lange termijn van 'de webcollectie Nederland' staat daarbij centraal.³ Daarnaast buigt de expertgroep zich over een nationaal programma gericht op praktische samenwerking, waarbij richtlijnen, tools en expertise gedeeld worden. Denk hierbij aan het delen van selectielijsten, harvest-software of opslagcapaciteit en het bespreken en adresseren van juridische kwesties. Niet iedere organisatie hoeft immers alles zelf te doen.

Studiedagen webarchivering

Ter ondersteuning van de samenwerking en de kennisdeling vinden er ook themabijeenkomsten plaats. Op 17 november 2016 kwamen zo'n 120 geïnteresseerden naar Beeld en Geluid om te praten en te leren over webarchivering tijdens de studiedag 'Een web van webarchieven'⁴, georganiseerd door de NCDD en het NDE. Het centrale thema van de dag was samenwerking. Deze dag was een vervolg van de studiedag in 2014 over webarchivering. De expertgroep zal ook initiatieven nemen om nieuwe studiedagen en workshops te organiseren.

Meedoen

Al deze initiatieven dragen bij aan een Nederlands web van webarchieven, om het heden te bewaren voor de toekomst.

Kennisnetwerk Archieven

Als samenwerkingsruimte is een Kennisplatform ingericht op de Pleio-omgeving van het Kennisnetwerk Archieven. Met een simpel aanmeldingsbericht kan iedere belangstellende lid worden, deelnemen aan discussies, bestanden delen en raadplegen. Praat mee over het archiveren van websites en sociale media in Nederland. Ga daarvoor naar pleio.nl (specifiek <https://goo.gl/zvgGst>).

We nodigen alle geïnteresseerde organisaties uit om hieraan deel te nemen, zodat we met elkaar de volgende stappen kunnen zetten. Alleen op die manier voorkomen we een digitaal gat in ons geheugen, omdat de geschiedenis van vandaag vooral online wordt geschreven. ■

Noten

1. Zie voor een lijst van websites opgenomen in het webarchief van de Koninklijke Bibliotheek: <https://goo.gl/GV2T0e>.
2. Zie hiervoor de resultaten van het inspectieonderzoek dat de Erfgoedinspectie in 2015 en 2016 heeft uitgevoerd: <https://goo.gl/LeJRr6>.
3. Informatie over de expertgroep is te vinden op: <https://goo.gl/MPv8Bz>.
4. Zie: <https://goo.gl/XIWuQ2>.

Met dank aan Carolien van Zuilekom (communicatieadviseur Nationale Coalitie Digitale Duurzaamheid) en Kees Teszelszky (onderzoeker webarchivering bij de Koninklijke Bibliotheek).

Marcel Ras ■ programmamanager bij de Nationale Coalitie Digitale Duurzaamheid.

Arnoud Goos ■ projectleider Netwerk Digitaal Erfgoed en coördinator Instroom bij het Nederlands Instituut voor Beeld en Geluid.

Workshop webarchivering

In het kader van de Week van het Digitaal Erfgoed organiseert de NCDD op woensdag 12 april 2017 een workshop webarchivering in het Dordrechts Museum. Meer hierover op <http://weekvanhetdigitaal erfgoed.nl>.