



Het coronacomplot is óók **digitaal erfgoed**

Hoe archiveer je coronamemes en complotsites?

Complotten bewaren voor later

WEBARCHIVERING

Coronajaar 2020 is ook op het internet een bijzonder jaar, maar weten we in de toekomst nog wel hoe dat eruitzag? De Koninklijke Bibliotheek probeert te voorkomen dat online coronacultuur verdwijnt en ja, ook complotsites.

Tekst **Menno van den Bos**
Illustratie **Timber Sommerdijk**



Voor Viruswaan.nl was het nèt te laat: voordat de oorspronkelijke site van de omstreden coronascepticus Willem Engel gearchiveerd kon worden, was-ie al vervangen door Viruswaarheid.nl. En zo zijn meer websites en -pagina's uit de eerste weken van de coronacrisis inmiddels niet meer te vinden, van RIVM-berichten tot het digitale kinderboek *Sem blijft thuis*. „Websites glippen soms door onze vingers”, vertelt Kees Teszelszky. „Wat op internet staat is niet eeuwig: de gemiddelde pagina staat maar drie maanden online.”

Teszelszky is conservator digitale collecties bij de Koninklijke Bibliotheek (KB) in Den Haag en legt samen met Peter de Bode, collectiespecialist bij de KB, een digitale coronacollectie aan. Het is een bonte verzameling online uitingen van leven met corona. Van Dagboekcorona.nl en Coronagedicht.nl, waar lockdown-emoities op rijm worden gezet (*Voorbij nu het schichtig begroeten/ we mogen elkaar weer ontmoeten*), tot complottheorieën. Samen met buitenlandse webarchieven werkt het duo ook aan een internationale collectie.

Doel: toekomstige generaties deze tijd laten begrijpen, vertellen De Bode en Teszelszky bij laatstgenoemde thuis aan de eettafel. Op tafel staat een scherm dat Teszelszky op zijn laptop aansluit. Zo kan hij collectie-items laten zien zonder schending van de anderhalvemeterregel. Dat woord staat trouwens in het online coronawoordenboek, een project van de website Taalbank.nl, eveneens in de collectie opgenomen. „Zo’n bron is als een gidsfossiel, waarmee archeologen aardlagen dateren. Op dezelfde manier kun je aan de opkomst van nieuwe woorden de veranderingen in een bepaalde tijd herkennen.”

Tijdreizigers

De KB bewaart sinds 2007 zogeheten digitaal geboren erfgoed, verdeeld over speciale collecties rond thema's als lhbt-

cultuur of het verdwijnende kloosterleven. De teller staat op 19.500 gearchiveerde websites. Dat is slechts een fractie van de tien miljoen sites die in Nederland geregistreerd zijn, dus de verzameling moet zo representatief mogelijk zijn voor de Nederlandse cultuur. „We proberen na te denken namens de historici die na ons zullen komen: wat hebben zij straks nodig om deze tijd te begrijpen?”, zegt Teszelszky, zelf gepromoveerd cultuurhistoricus. „We zijn een soort tijdreizigers.”

Op een ochtend in februari keken De Bode en hij elkaar aan: moesten ze niet ook iets met dat corona? Niet lang daarna barstte de pandemie los en waren ze „24 uur per dag” met de coronacollectie bezig. Als ze iets interessants zien, moeten ze snel handelen, omdat ze wettelijk verplicht zijn om de website-eigenaar per brief vier weken te geven om opname in het archief te weigeren - vier weken waarin een site met wat pech offline kan gaan.

Teszelszky en De Bode zijn de motor van het tienkoppige team Webarchivering (drie fte), een aparte afdeling bij de KB. Hun vak staat nog in de kinderschoenen. Teszelszky was in 2016 pas de eerste in Nederland die speciaal werd aangeworven om digitaal geboren erfgoed te beheeren - eerder werkte hij onder meer als webmaster, materieel-erfgoedbeheerder en universitair onderzoeker van middel-eeuwse en Hongaarse geschiedenis. „In mijn eerste jaar bij de KB heb ik mezelf opgeleid. Ook in het buitenland zijn er weinig mensen met mijn functie.”

De hoofdtaak van de nationale bibliotheek is alle boeken, tijdschriften en kranten bewaren die in Nederland verschijnen, waarvan een deel wordt gedigitaliseerd. Hebben die historici van de toekomst aan alle kranten die nu vol coronanieuws staan eigenlijk niet genoeg? Nee, zegt het duo, want daarin mist het aanzienlijke deel van ons dagelijks leven dat zich online afspeelt. „Een krant schrijft misschien dat er een nieuwe site is, maar

niet hoe mensen die site invulden.”

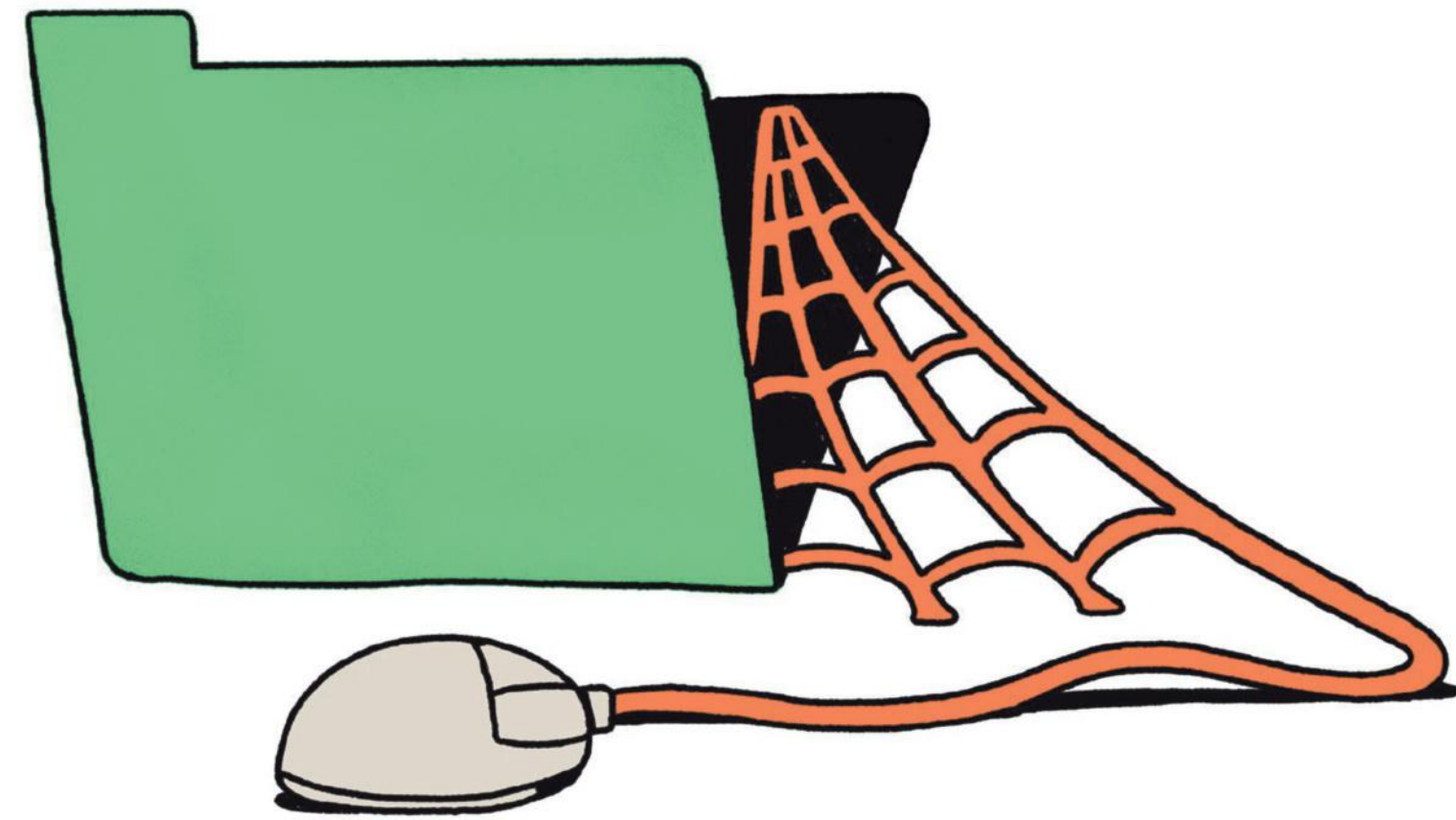
Het belang van digitaal archiveren groeit ook omdat die van het papier krimpt: zo houden steeds minder mensen een dagboek bij. Teszelszky: „Toen ik een boek schreef over een 16de-eeuwse Hongaarse diplomaat, kon ik zijn tochten van dag tot dag reconstrueren via zijn brieven, aantekeningen en inscripties in gastenboeken van vrienden die hij bezocht. Datzelfde doen voor een diplomaat uit onze tijd is een stuk lastiger. Die plaatst misschien een update op Facebook of Twitter, maar die berichten worden vaak na een tijdje verwijderd of verdwijnen in de databases van die techbedrijven, onbereikbaar voor onderzoekers.”

Online sporen vervagen, kortom, sneller dan inkt op papier. „Van veel middeleeuwse kronieken heeft de KB vijf tot tien reserve-exemplaren. Maar als een homepage uit de jaren negentig offline gaat, is-ie voorgoed verloren.”

Staatscollectie

De KB'ers werken met een crawler, een computerprogramma dat zich door een website heenwerkt en onderweg alle bouwsteentjes lospeutert: tekst, afbeeldingen, html-code, stylesheets. „Het is als een set lego. Als je de gearchiveerde site oproept, wordt uit al die losse elementjes de website gereconstrueerd.” Er zijn maar drie mensen ter wereld die de techniek van deze crawler volledig begrijpen, stelt Teszelszky. „We waren een keer op een conferentie waar ze alledrie waren, en toen viel er eentje van de trap. Dan denk je gelijk: *shit*.”

Worden veel sites niet bewaard door het Amerikaanse Internet Archive, bekend van de Wayback Machine (een soort Google voor oude websites)? „We horen vaak: daar zit alles toch al in? Maar dat is niet zo. 90 procent van de Nederlandse homepages uit de jaren negentig die nog online staan ontbreken bijvoorbeeld. En van veel websites bewaren ze alleen de homepage, waardoor je niet kunt door-



Sites van coronasceptici gaan sneller offline dan gemiddeld

Hoe doen zij het?

De Britten halen alles binnen

De Nederlandse wet schrijft voor dat elke website-eigenaar opname in de collectie van de KB mag weigeren. Liever zouden Kees Teszelszky en Peter de Bode in één keer alle Nederlandse sites kunnen opsorpen, zoals hun collega's bij The British Library. „Daar gaan ze jaarlijks met taart rond een grote knop staan, waarna alles wordt binnengehaald en ze weer iets anders kunnen gaan doen”, zegt Teszelszky. Groot-Brittannië heeft, net als onder andere Frankrijk, een zogeheten wettelijk depot-regeling waardoor website-eigenaren geen opt-out-brief gestuurd hoeft te worden. De KB'ers willen zo'n regeling ook in Nederland: het zou automatisch de redding betekenen van talloze antieke homepages. Die zijn nu bijna onvindbaar omdat geen enkele andere site er nog naar linkt en vice versa. Het zijn eilandjes in een uitgestrekte oceaan, die niet meer op de radar van Google verschijnen en die de webarcheologen alleen kunnen vinden via een analoge aanwijzing, zoals een surftip in een vergeeld computerblad. Een bewerkelijke werkwijze waardoor ze niet op kunnen tegen de snelheid waarmee het web door de vergetelheid wordt opgeslokt.

klikken.” De Bode: „Of het kan wel maar dan kom je bij een latere versie van die website terecht. Voor wetenschappers zijn dat onbetrouwbare bronnen.”

In de covidcollectie zitten nu 310 websites, inclusief sites die nog in de wacht staan. Een belangrijk onderdeel zijn plekken waar kritiek op het kabinetsbeleid wordt geleverd, zoals de opinieblogs *Joop*, *ThePostOnline* en *GeenStijl*. Ook bij Maurice de Hond belandde een brief op de mat. De Bode: „We proberen alle meningen over corona vast te leggen, zonder oordeel.”

Zelfs websites van complotdenkers krijgen een plekje, zoals corona-nuchtheid.nl. ‘Coronasceptici’ vormen een spiegel van onze post-truth-tijd, zegt Teszelszky, waarin mensen op elke vorm van autoriteit kritisch zijn. „Vroeger ging je naar de huisarts en nam je wat die zei voor waar aan. Nu nemen patiënten een stapel eigen bronnen van internet mee. Door corona ontstaat zelfs in hoogopgeleide kringen discussie over medische feiten.”

Sites van coronasceptici moeten extra snel bewaard worden, omdat ze nog sneller offline gaan dan gemiddeld. Sommige eigenaren willen onder geen beding in de ‘staatscollectie’ worden opgenomen, zegt Teszelszky, anderen zijn juist vereerd, „omdat ze het idee hebben dat ze met hun verzet met iets historisch bezig zijn”.

Opname in het archief is echter geen stempel van goedkeuring, zegt De Bode: „Omdat we in de brief refereren aan het belang dat de Unesco aan digitaal erfgoed hecht, kopte GeenStijl ludiek met ‘Unesco approved’. Maar we keuren natuurlijk niet, we zijn een onafhankelijke instelling.” Wat weer iets anders is dan objectief. „Het blijft een intuïtieve selectie. We zijn daarom blij dat veel mensen ons tips sturen. Zo behouden we een divers beeld.”

Nepnieuws bewaren

Tot voor kort was het bij de nationale bibliotheek *not done* om aanstootgevende websites te archiveren. „In het archief zat lang een bovengemiddeld aantal websites van musea of over literatuur, maar dat geeft natuurlijk geen representatief beeld van internetcultuur”, zegt Teszelszky - veel te braaf. Complotten en nepnieuws zijn het andere uiterste: moet je dat wel willen bewaren? „De wetenschappers van de toekomst zullen weer op andere manier kijken naar wat waar en niet waar is. Het oordeel daarover laten we aan hen.”

Veel coronacultuur ziet het licht op sociale media - helaas mogen de KB'ers van wat mensen daar delen wettelijk niets in de collectie opnemen. Maar daar heeft Teszelszky iets op bedacht: een officieuze privécollectie met screenshots van socialmediaberichten en memes over corona. Zoals Instagramfoto's van stoepkrijttekeningen van kinderen die tijdens de eerste lockdown hun buurt een hart onder de riem staken. „In het begin lag ik veel wakker door alle coronaperikelen. Toen lag ik de hele nacht op social media te scrollen en screenshots te maken. Nu denk ik: wauw, wat ben ik blij dat ik dat toen heb gedaan.”