

NL-blogosfeer

Speciale webcollectie Koninklijke Bibliotheek

Versie 25 juni 2018

Laatste kleine wijzigingen aangebracht op: 23-02-2021



Colofon

Uitgever:
Koninklijke Bibliotheek (KB), 2021

Redactie:
Peter de Bode en Kees Teszelszky

Beeld:
Word Pictogrammen

Met dank aan:
Anne Helmond en Esther Weltevrede

Meer informatie:
webarchivering@kb.nl

© Koninklijke Bibliotheek, Den Haag.

Inhoud

1	<i>Inleiding</i>	4
1.1	Wat is de webcollectie van de Koninklijke Bibliotheek?	4
1.2	Waarom is deze speciale webcollectie opgezet?	4
1.3	Waarom archiveert de Koninklijke Bibliotheek weblogs?	6
2	<i>Historische context van de webcollectie</i>	8
2.1	Geschiedenis van het weblog	8
2.2	Term trends	11
3	<i>Digitaal erfgoed</i>	13
3.1	Wat is de waarde van de collectie als digitaal erfgoed?	13
4	<i>Onderzoeksmethode voor selectie</i>	14
5	<i>Selectiemethode</i>	17
5.1	Selectiecriteria voor archivering (Koninklijke Bibliotheek)	17
5.2	Selectiecriteria voor NL-Blogosfeer collectie	19
5.3	Excel uitleg: bronbestanden / geselecteerd via	20
6	<i>Contactadres, analoge contacten</i>	24
6.1	Contactadres / verweesde websites	24
7	<i>Techniek en technische problemen</i>	25
7.1	WCT / Heritrix	25
7.2	Technische problemen	25
7.3	Gebruikersproblemen	26
8	<i>Geselecteerde sites en beschrijving topstukken</i>	28
8.1	Kolommen	28
8.2	Formats & indeling personen	32
8.3	Genre	35
8.4	Parels & beschrijving	40
9	<i>Linkanalyse</i>	43
9.1	Hyphe linkanalyse	43
10	<i>Datasets</i>	48
10.1	Helmond lijst (bonuslijst Helmond Dutch blogosphere schoon en ontdebelt)	48
10.2	Ongeldige links (bonuslijst ongeldige links)	48

10.3	KB blogs (bonuslijst KB blogs)	49
11	<i>Status geselecteerde websites per 31-08-2020</i>	50
12	<i>Conclusie</i>	51
12.1	Algemene conclusie	51
12.2	Aanbevelingen	53
13	<i>English summary</i>	55
14	<i>Verder onderzoek</i>	56
14.1	Waar zouden nog meer blogs gevonden kunnen worden?	56
14.2	Onderzoeksideeën	56
15	<i>Literatuur</i>	58
15.1	Literatuur	58
15.2	Websites	59
15.3	Artikelen op websites (met Internet Archive Wayback link waar mogelijk).	60
16	<i>Artikelen gepubliceerd naar aanleiding van deze collectie</i>	63
16.1	Wetenschappelijk	63
16.2	Populair	63
16.3	NL-blogosfeer wordt onderdeel van het blog	63
17	<i>Bijlage 1: Internetarcheologie</i>	65
17.1	Het belang van internetarcheologie	65
17.2	Internetarcheologie wishlist	67
18	<i>Bijlage 2: Definitie uit gelezen literatuur</i>	69
19	<i>Bijlage 3: Thema's / genres: welke categorieën houden anderen aan?</i>	72
20	<i>Bijlage 4: Lifestyle VS Lifelog analyse</i>	76

1 Inleiding

1.1 Wat is de webcollectie van de Koninklijke Bibliotheek?

De KB-webcollectie is een unieke collectie van digitaal geboren bronmateriaal en data over de Nederlandse digitale cultuur, die nergens anders ter wereld in deze vorm wordt verzameld, bewaard en kan worden geraadpleegd. Bovendien is dit materiaal binnen afzienbare tijd niet meer op het web te vinden, maar alleen nog maar bij de KB te bestuderen. Binnen de KB webcollectie wordt gestreefd representatieve websites uit Nederland of over Nederland te selecteren, waarbij de primaire focus gericht is op websites (waarbij tekst overheersend is) die betrekking hebben op Nederlandse taal, cultuur, politiek, geschiedenis, samenleving en de informatiestroom van de Nederlandse overheid. De selectie van de afgelopen jaren is ook afhankelijk geweest van wat op het moment van archivering technisch en organisatorisch mogelijk was. De KB is in 2007 begonnen met het verzamelen van websites. Hierdoor heeft de KB gaten in de webcollectie over de beginperiode van het internet tussen 1993 en 2000 en de periode 2001-2007, die door middel van retrocollectionering wordt opgevuld.

De huidige Auteurswet biedt weinig ruimte voor vrije toegang tot de webcollectie verzameld in het KB-webarchief. Iedereen kan de webcollectie binnen de muren van de KB bestuderen (on site), maar de content mag niet worden gekopieerd (tenzij voor privégebruik of studie) of openbaar worden gemaakt. Een gebruiker kan niet op afstand een inschatting doen van het nut van de collectie voor zijn of haar onderzoek. Hierdoor is het bieden van context bij deze webcollectie net zo belangrijk als het beschikbaarstellen van de inhoud.

De KB streeft ernaar om de webcollecties aan te bieden met zoveel mogelijk achtergrondinformatie. De webcollectie krijgt zo een context met redactionele informatie. Hiermee wordt helderheid geboden over de historische context van de collectie, hoe de webcollectie is opgebouwd en wat de wensen zijn voor de toekomst. Daarnaast wordt uiteengezet hoe de primaire gebruikers, wetenschappers, optimaal gebruik kunnen maken van de webcollectie.

1.2 Waarom is deze speciale webcollectie opgezet?

De afgelopen drie maanden heb ik (Iris Geldermans) als stagiaire voor de Koninklijke Bibliotheek gewerkt aan het vormen van de collectie Nederlandse weblogs voor het webarchief: NL-blogosfeer.¹ Ik heb tijdens deze stageperiode ontdekt dat weblogs belangrijk primair bronmateriaal zijn voor historisch onderzoek omdat zij de gebeurtenis in de context van de tijd plaatsen. Bloggers houden zich continue bezig met de actualiteit. Veel blogs zijn een *bottom-up* bron: zij geven een kijkje in het dagelijks leven en de meningen van het volk. Weblogs laten zien hoe mensen reageerden op actuele thema's zoals verkiezingsuitslagen, festivals, Pim Fortuyn, #MeToo, veranderende modetrends, de reactie op 9/11, MH17 en de vuurwerkcramp in Enschede. De blogs zijn echter ook kwetsbaar. Zoals zoveel websites verdwijnen weblogs van het internet. Hierbij verdwijnen ook archieven die jaren teruglopen en de bijbehorende reacties en discussies op historische gebeurtenissen.

¹ De rest van het verslag: wanneer 'ik' het over 'ik, mijzelf, mij' heb dan gaat het om de stagiaire webarchivering / Iris Geldermans.

Een goed voorbeeld van het verdwijnen van weblogs zijn de gebeurtenissen omtrent het weblog.nl domein en aanbieder Sanoma.² Op 5 april 2003 begon de blogdienst web-log.nl met zijn dienst als blogplatform. Gebruikers konden met dit weblogstelsel makkelijk hun eigen blogs bijhouden. Er was op dat moment geen soortgelijk platform in Nederland. Er waren alleen Amerikaanse equivalenten die al langer bestonden zoals Blogger.³ Na één jaar werd het bedrijf opgekocht door Ilse Media (onderdeel van Sanoma). In 2005 verwelkomde het platform zijn honderdduizendste blog.⁴

In 2011 begon het mediabedrijf Sanoma met verandering van het web-log domein. Het idee was simpel: verwijder het tussenliggende streepje waardoor de doorzoekbaarheid van het domein vergroot zou worden. Daarnaast stopte TypePad met het aanbieden van zijn service waardoor overgestapt moest worden naar WordPress. Het werd een debacle: eind augustus ging de blogdienst offline waardoor 400.000 blogs opeens offline gingen.⁵ Extra vervelend hierbij was het vergaan van het Volkskrantblog.

De Volkskrant was in 2005 begonnen met het aanbieden van een blogdienst voor lezers: het Volkskrantblog (vkblog.nl) was een mengeling van redacteurs en lezers waarbij iedereen (ook mensen die geen abonnee waren) een bijdrage kon leveren.⁶ In 2011 kwamen de Volkskrant en Sanoma overeen dat Weblog.nl de VKblogs zou gaan hosten. De genoemde 'volksbloggers' zouden hun Volkskrantblog kunnen importeren in Weblog.nl.⁷

Op 1 december 2012 verwijderde Sanoma alle inactieve blogs (een weblog waarop al langer dan twaalf maanden niet was ingelogd) van het domein weblog.nl die zij konden vinden. De blogs zouden de database traag maken en actieve webloggers hinderen.⁸ Januari 2013 kondigde Sanoma aan dat VK blog niet naar WordPress omgezet zou worden.⁹ De VK bloggers zouden hun blogs niet terugkrijgen. Op 1 maart 2013 sloot weblog.nl definitief zijn deuren. Gebruikers moesten hun blogs

² Zie als voorbeeld het Venturo weblog in de Wayback Machine van het Internet Archive,,
<https://web.archive.org/web/20130508153845/http://venturo.weblog.nl/> (16-05-2018).

³ <https://web.archive.org/web/20130416072539/http://incredibleadventure.nl/2013/01/sanoma-stopt-met-weblog-nl-een-terugblik-met-de-oprichter/> (16-05-2018).

⁴ <http://webwereld.nl/e-commerce/26807-web-log-nl-begroet-100-000ste-weblog> (16-05-2018).
<https://web.archive.org/web/20180524114958/http://webwereld.nl/e-commerce/26807-web-log-nl-begroet-100-000ste-weblog> (24-05-2018).

⁵ <http://webwereld.nl/e-commerce/231-de-8-maanden-migratiemalaise-van-web-log-nl> (16-05-2018).
<https://web.archive.org/web/20180524115304/http://webwereld.nl/e-commerce/231-de-8-maanden-migratiemalaise-van-web-log-nl> (24-05-2018).

⁶ <http://webwereld.nl/e-commerce/28258-volkskrant-geeft-lezers-een-stem-via-nieuw-blog> (16-05-2018).
<https://web.archive.org/web/20180524115401/http://webwereld.nl/e-commerce/28258-volkskrant-geeft-lezers-een-stem-via-nieuw-blog> (24-05-2018).

⁷ <http://webwereld.nl/development/57010-volkskrant-blogs-al-maanden-offline-door-web-log-nl> (16-05-2018).
<https://web.archive.org/web/20180524115459/http://webwereld.nl/development/57010-volkskrant-blogs-al-maanden-offline-door-web-log-nl> (24-05-2018).

⁸ <http://webwereld.nl/e-commerce/58525-web-log-nl-deletet-oude-blogs> (09-05-2018).
<https://web.archive.org/web/20180524115551/http://webwereld.nl/e-commerce/58525-web-log-nl-deletet-oude-blogs> (24-05-2018).

⁹ <http://webwereld.nl/e-commerce/59033-vkblogs-sneuvelen-in-nieuwste-web-log-nl-migratie> (16-05-2018).
<https://web.archive.org/web/20180524115657/http://webwereld.nl/e-commerce/59033-vkblogs-sneuvelen-in-nieuwste-web-log-nl-migratie> (24-05-2018).

handmatig exporteren als zij wilden dat hij bewaard bleef.¹⁰ Ondanks dat een patchwork en incomplete versie van het VKblog nog bewaard wordt in het Internet Archive of op een website van iemand die zijn archief op tijd heeft gered, is het merendeel van deze web-log en 'volksbloggers' gemeenschap nu voor altijd verloren.¹¹

Een ander voorbeeld van een verloren blogparel is Loglands. In 2004 besloot een groep van vier bloggers op Lowlands een liveblog bij te houden via hun mobieltjes. Fans van het festival weten dat dit het jaar is waar de vermiste 'Theo' geïntroduceerd werd als (een tot op het heden voortdurende) lopende grap van het festival.¹² 'Raak niet in de veronderstelling dat je Theo of Michael ooit zult vinden.'¹³ Dit (eenmalige) liveblog is net als VKblog alleen nog te vinden in het Internet Archive. Maar is ook, net als VKblog, een lappendeken van snapshots waarbij de voorpagina voor het laatst is gearchiveerd op 11 augustus 2007 en de archiefpagina in februari 2006.¹⁴

Om te voorkomen dat nog meer van deze Nederlandse blogcultuur verloren gaat, heeft de Koninklijke Bibliotheek besloten een collectie blogs aan te leggen onder de naam: NL-blogsfeer. De NL-blogsfeer collectie heeft als doel een representatief beeld geven van de Nederlandse blogcultuur. Daarnaast kunnen websites ook opgenomen worden vanwege hun nationaal cultureel historisch belang.

1.3 Waarom archiveert de Koninklijke Bibliotheek weblogs?

Directe aanleiding voor het vormen van de collectie was het onderzoek *Where do bloggers blog? Platform transitions within the historical Dutch blogosphere* door Esther Weltevrede en Anne Helmond.¹⁵ In dit onderzoek analyseren Weltevrede en Helmond de Dutch Blogosphere aan de hand

¹⁰ <https://tweakers.net/nieuws/86711/gratis-weblogdienst-weblog-punt-nl-sluit-deuren.html> (09-05-2018).
<https://web.archive.org/web/20180524115821/https://tweakers.net/nieuws/86711/gratis-weblogdienst-weblog-punt-nl-sluit-deuren.html> (24-05-2018).

¹¹ <http://www.duurzamebrabanders.nl/blog/tag/vkblog/> (bekeken op 11-06-2018) Henk Daalder was één van de gelukkigen die op tijd zijn VKblog back-upte naar een zelfstandige website. De posts die hier staan gaan over het VKblog. Of de rest van zijn archief online staat of alleen op CD is onduidelijk: <http://www.duurzamebrabanders.nl/blog/2011/01/maak-een-mooie-backup-op-een-gemakkelijke-manier/>
<https://web.archive.org/web/20180611093652/http://www.duurzamebrabanders.nl/blog/2011/01/maak-een-mooie-backup-op-een-gemakkelijke-manier/>
<http://www.duurzamebrabanders.nl/blog/2011/01/zijn-er-nog-duurzame-bloggers-bij-het-vkblog/>
<https://web.archive.org/web/20180611093442/http://www.duurzamebrabanders.nl/blog/2011/01/zijn-er-nog-duurzame-bloggers-bij-het-vkblog/> Het blog is in juni 2018 aangemerkt door mij voor archivering.

¹² Meeuwssen, F., *Bloghelden* (2010) 64.

¹³ <https://www.yellowtipi.nl/blog/lowlands-voor-beginners> (16-05-2018).
<https://web.archive.org/web/20180524120049/https://www.yellowtipi.nl/blog/lowlands-voor-beginners> (24-05-2018).

¹⁴ https://web.archive.org/web/20060227215040/http://www.loglands.nl:80/archives/archive_2004-m08.php
<https://web.archive.org/web/20070811060058/http://www.loglands.nl> (16-05-2018).

¹⁵ Weltevrede, E. and Helmond A. (2012) Where do bloggers blog? Platform transitions within the historical Dutch blogosphere. *First Monday* 17(2). DOI: [10.5210/fm.v17i2.3775](https://doi.org/10.5210/fm.v17i2.3775).
Te vinden op: <http://firstmonday.org/ojs/index.php/fm/article/view/3775> (31-05-2018). Het hele artikel staat hier genoemd onder het adres <http://journals.uic.edu/ojs/index.php/fm/article/view/3775/3142>.
<https://web.archive.org/web/20180531140657/http://journals.uic.edu/ojs/index.php/fm/article/view/3775/3142> (31-05-2018)

van expertlijsten gevonden door onderzoek in literatuur (zoals *Bloghelden* van Frank Meeuwsen¹⁶) en interviews met bloggers die onder andere betrokken waren bij de Dutch Bloggie awards.

De Koninklijke Bibliotheek heeft besloten naar aanleiding van dat onderzoek een stagiaire in huis te nemen die kritisch zou gaan kijken naar de selectie en reconstructie van Helmond en Weltevrede en vervolgens de collectie zou uitbreiden, analyseren, beschrijven en archiveren. De collectie weblogs is bedoeld als primaire bron van of reflectie op de digitale cultuur in Nederland.

¹⁶ Bekijk hier de website van het boek en download de PDF van het boek: <http://bloghelden.nl/>.
<https://web.archive.org/web/20180531141015/http://bloghelden.nl/>
<https://web.archive.org/web/20180531141209/https://raw.githubusercontent.com/frankmeeuwsen/Bloghelden/master/bloghelden-boekversie.pdf> (31-05-2018).

2 Historische context van de webcollectie

2.1 Geschiedenis van het weblog

Voorlopers van het blog

Voor er weblogs waren, waren er mailing lists, bulletin boards en de Digitale Stad.¹⁷ Deze drie voorlopers werden gebruikt om via het internet gemeenschappen op te richten, discussies te houden en actuele nieuwtjes te verspreiden.

Neem bijvoorbeeld de definitie van een 'post'. Tegenwoordig is een post een berichtje op een blog of social media toepassing. In 1997 werd een 'post' omschreven als: 'Bericht dat iemand in een nieuwsgroep plaatst. Iedereen die de nieuwsgroep leest kan het bericht lezen.'¹⁸ In 2000 is de definitie alweer enigszins veranderd: 'Een artikel dat is gepubliceerd op of verstuurd naar een mailinglijst of Usenet-nieuwsgroep.'¹⁹ Een maillijst kreeg de volgende beschrijving in 1997: 'Speciaal e-mailadres dat alle binnenkomende post doorstuurt naar een lijst met abonnees. Elke mailinglijst is gewijd aan een specifiek onderwerp, dus abonneer je alleen op lijsten die voor jou interessant zijn.'²⁰ Een definitie uit 2000 van de maillist toont dat het ook een belangrijk element bevat waar weblogs bekend om staan: mensen kunnen reageren op verstuurd berichten en er ontstaat hierdoor een levendige discussie: 'op de lijst staan speciale eigen e-mailadressen, en alles wat naar een persoon op die lijst wordt gestuurd, komt bij iedereen op de lijst terecht. Omdat mensen ook weer antwoorden op dat bericht is het resultaat een levendige conversatie.'²¹

Deze voorlopers werden gevolgd door / overlappen met de opkomst van weblogs en blogs die de KB wil archiveren en waar deze collectie over gaat. Na de weblogs en blogs op internetpagina's volgt de social media en het microbloggen. Social media websites en apps zijn op dit moment nog niet te archiveren en worden daarom niet verzameld door de KB.

Het (we)blog

Het weblog zelf is door de jaren heen van vorm veranderd. Er zijn hierbij twee algemene ontwikkelingen te benoemen. Ten eerste veranderde het blog van mediavorm of format. In het begin waren blogs alleen tekst met ondersteunende afbeeldingen of een blog met uitsluitend afbeeldingen (of uitsluitend tekst). Later, toen de techniek het toeliet, ontstonden er ook blogs met audiofragmenten (vanaf 2004) of zelfs de gehele blog in audioformat: de podcast. Tenslotte volgden de

¹⁷ Voor meer informatie over bulletin boards zie Internationaal Instituut voor Sociale Geschiedenis. <https://socialhistory.org/nl/collections/occasio> en <https://socialhistory.org/sites/default/files/docs/collections/occasio-background.pdf> Voor meer informatie over de digitale stad kijk op de website van het Amsterdam Museum: <https://hart.amsterdam/nl/page/521/re-dds>

¹⁸ Levine, J.R., C. Baroudi en M. Levine Young, *Internet voor dummies. Dé referentie voor iedereen!* (Amsterdam 1997) 317

¹⁹ Levine, J.R., C. Baroudi en M. Levine Young, *Internet voor dummies. Dé referentie voor iedereen!* (Amsterdam 2000) 347

²⁰ Levine, Baroudi en Levine Young, *Internet voor dummies. Dé referentie voor iedereen!* (1997) 323.

²¹ Levine, Baroudi en Levine Young, *Internet voor dummies. Dé referentie voor iedereen!* (2000) 219 – 220.

videofragmenten en vlogs (blogs die met name videofragmenten bevatten of uitsluitend videofragmenten zijn).

Een tweede algemene ontwikkeling was dat het blog in het vroege stadium echt nog een persoonlijk dagboek was. De originele term 'weblog' stond voor linkdumps: surfers die een dagboek van links bijhielden die zij vonden tijdens hun zwerftochten over het internet. Tijdens de opkomst van het internet hielden deze pioniers zich bezig met het in kaart brengen van het internet voor nieuwe gebruikers. De bloggers wisten de laatste nieuwtjes en de leukste websites aan te wijzen. Het blog groeide daarna al snel uit tot een persoonlijk dagboek met teksten en foto's. Snel daarop werd het blog nog breder: zakelijke lieden en politici gingen er gebruik van maken. Vanaf 2004 gingen bijvoorbeeld Amerikaanse politici blogs gebruiken om hun mening te verspreiden en stemmen te trekken. Het is vanaf dat moment dat het blog als uitsluitend persoonlijk dagboek verandert naar een multifunctioneel middel waarbij alle mogelijke onderwerpen aan bod komen.²² Tegenwoordig kan een blog dan ook van alles zijn: het kan een dagboek zijn over het leven van een persoon, een financieel blog, een shocklog (zoals GeenStijl), entertainment of lifestyle.

Historische definities:

2001: 'Bloggers zijn mensen die een "Weblog", een soort dagboek, van hun surftochten over het Web bijhouden en publiceren.'²³

2005 Van Dale: 'Blog (...) persoonlijk dagboek op internet = blogsitem, weblog; - ook als tweede lid in samenst. Als de volgende, waarin het eerste lid het thema, onderwerp van het blog noemt: kunstblog, literatueurblog, poëzieblog, wiskundeblog'.²⁴ Op dezelfde pagina worden ook de volgende termen uitgelegd: blogbericht, blogboek, bloggen, blogger, bloghoppen, blogomie ('economische bedrijvigheid die verband houdt en voortvloeit uit communicatie via weblogs'²⁵), blogosfeer, blogpost, blogposting, blogsite, blogstek (synoniem voor weblog) en blogware.²⁶

2005: 'Een blog is een website die bestaat uit een reeks afzonderlijke berichten, die ook wel posts of postings genoemd worden. Deze posts of bijdragen zijn vrijwel altijd voorzien van een datum en tijd en de meest recente bijdrage staat bovenaan het lijstje. Een echte weblog wordt regelmatig bijgewerkt, vaak zelfs dagelijks. Bezoekers van een weblog kunnen hun reacties kwijt op een blog. Zo ontstaan vaak interessante discussies – een fenomeen waardoor blogs zo populair zijn geworden en al lang niet meer alleen thuishoren onder het hoofdstuk "internetamusement".'²⁷

²² Meindersma, C., *Wetboek voor bloggers. De (on)geschreven regels voor bloggers, journalisten en content creators* (Den Haag 2016) 28.

²³ Hofstede, M., *Het zoekboek voor het web* (Rijswijk 2001) 31.

²⁴ In 1999 stond er in de Van Dale nog niets over weblogs. Dit veranderd in de 2005 editie. Boon, T den, Hendrickx, R. en N. van der Sijs, *Van Dale. Groot woordenboek van de Nederlandse taal* (Utrecht/Antwerpen 2005) 520.

²⁵ Den Boon, Hendrickx.en Van der Sijs, *Van Dale*. (2005) 520.

²⁶ Den Boon, Hendrickx.en Van der Sijs, *Van Dale*. (2005) 520.

²⁷ Holthuijsen, D., *Weblogs. Uw persoonlijke blog op het web!* (Haarlem 2005) 10.

2005 en 2007: 'Afkorting weblog, een persoonlijk dagboek op internet. Elke gek kan een blog bijhouden, en veel gekken doen het dan ook.'²⁸

2013 (en 2018) Van Dale:

'blog (...) weblog. Blogger (...) iem. Die een weblog bijhoudt'.²⁹

'Weblog (...) actuele website waarop regelmatig korte stukjes, foto's, filmpjes enz. verschijnen, al dan niet uit de persoonlijke sfeer'.³⁰

2015 (en 2018): 'Een weblog, meestal afgekort tot blog, is een openbaar online dagboek waarin iemand min of meer regelmatig updates post. Een blog gebruikt software waarmee je eenvoudig berichtjes kunt posten met behulp van je webbrowsers; je hebt geen extra software nodig. (...) De meeste blogs worden regelmatig geüpdatet door één auteur en bevatten lijsten met redelijk recente berichten, vergelijkbaar met een online dagboek of agenda. De recentste berichten staan bovenaan. Sommige blogs zijn wat complexer en gaan over meerdere onderwerpen en bevatten naast (of in plaats van) woorden ook afbeeldingen, maar het idee van relatief korte berichten die vrij vaak worden geüpdatet blijft gehandhaafd. (...) Nu blogs populairder worden, hebben veel websites ook blogs om nieuws, roddels of persoonlijke verhalen te posten. (...) Sommige websites geven de laatste blogposts een prominente plek. Je kunt vaak klikken op een koppeling om oudere berichten te zien.'³¹

Tijdslijn:

1994: online dagboeken ontstaan, na de uitvinding van het http-protocol.³²

1 december 1997: eerste weblog in Nederland, Daily Wacko. Dit was een parodienieuwsbrief op de dagelijkse email van de Daily Planet (door Francisco van Jole in opdracht van internetprovider Planet Internet).³³

17 december 1997: de term weblog wordt geïntroduceerd door Jorn Barger: men hield namelijk een logboek bij van wat zij op het web vonden.³⁴

²⁸ Levine, J.R. en C. Baroudi en M. Levine Young, *Internet en e-mail voor dummies. Dé referentie voor iedereen!* (Amsterdam 2007) 328. Levine, J.R. en C. Baroudi en M. Levine Young, *Internet en e-mail voor dummies. Dé referentie voor iedereen!* (Amsterdam 2005) 356.

²⁹ Boer, T. de, *Van Dale Pocketwoordenboek Nederlands* (Utrecht 2013) 87.

³⁰ De Boer, *Van Dale Pocketwoordenboek Nederlands* (2013) 620. <http://www.vandale.nl/gratis-woordenboek/nederlands/betekenis/weblog> (16-04-2018). <https://web.archive.org/web/20180524123949/https://www.vandale.nl/gratis-woordenboek/nederlands/betekenis/weblog> (24-05-2018).

³¹ Levine, J.R. en M. Levine Young, *Internet voor dummies. Dé referentie voor iedereen!* (Amersfoort 2015) 303 - 304. Levine, J.R. en M. Levine Young, *Internet voor dummies. Dé referentie voor iedereen!* (Amersfoort 2018). 287 – 288.

³² <https://nl.wikipedia.org/wiki/Weblog> (09-04-2018). <https://web.archive.org/web/20180524123500/https://nl.wikipedia.org/wiki/Weblog> (24-05-2018).

³³ <https://nl.wikipedia.org/wiki/Weblog> (09-04-2018).

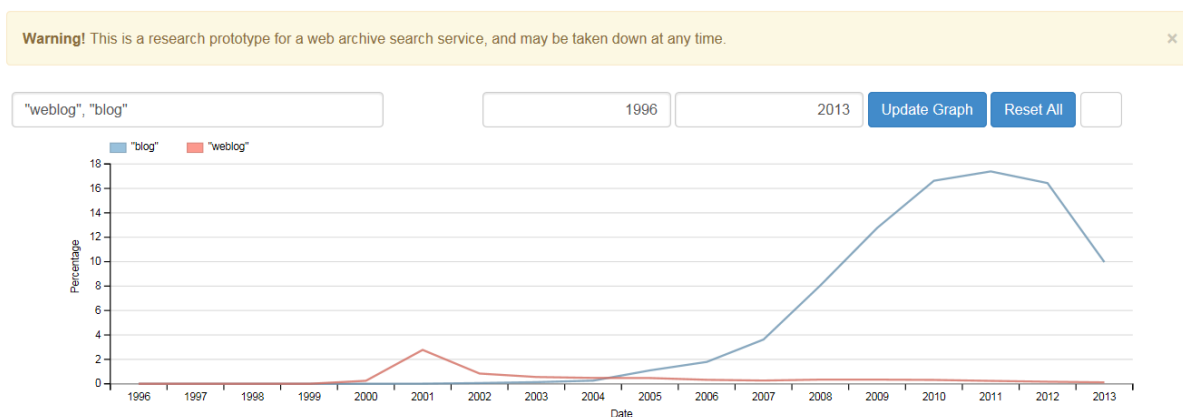
³⁴ <https://nl.wikipedia.org/wiki/Weblog> (09-04-2018).

Mei 1999: de term blog wordt bedacht door Peter Merholz (bij wijze van grap).³⁵

2004 – 2007/08: populariteit van postcasts (audio fragmenten) neemt toe tot in 2005. In 2005 wordt namelijk YouTube opgericht. Dan neemt de vlog het over. Tegelijkertijd ontstaat ook de microblog: in 2006 wordt bijvoorbeeld Twitter (max 140 tekens) opgericht en vanaf 2007 Tumblr.³⁶

2.2 Term trends

Shine, een applicatie waarmee het Britse webarchief doorzocht kan worden (van 1996 tot april 2013), laat nog eens concreet zien hoe vaak de termen 'blog' (blauw gekleurd) en 'weblog' (rood gekleurd) aanwezig zijn in hun webarchief.³⁷ Zoals je ziet is de term 'weblog' in gebruik van 1999 – 2003 om tijdens 2003 – 2005 plaats te maken voor de term blog. Die term wint dan ook van het begrip 'weblog'.



Een tweede interessante trendanalyse is uitvoerbaar via Google Trends. Google Trends geeft weer hoe vaak een zoekterm is gebruikt op google. Dit gebeurt niet in absolute cijfers maar in een gewogen nummer. De data wordt geïndexeerd tot maximaal 100: 100 is de maximale zoekinteresse voor een geselecteerde tijd en plaats.³⁸ In deze analyse is de periode 2004 – 2018 gebruikt (verder terug was niet mogelijk) voor specifiek Nederland.³⁹ In Nederland gaat men dus eind 2005, 2006, steeds minder zoeken op de term 'weblog' en steeds meer op de term 'blog'. Maar de maximale zoekinteresse heeft 'blog' nooit gehaald. Alleen 'weblog' haalde dat in augustus 2005. Op het moment

³⁵ <https://nl.wikipedia.org/wiki/Weblog> (09-04-2018).

³⁶ Meindersma, C., *Wetboek voor bloggers* (2016) 29.

³⁷ Shine is een prototype historische zoekmachine ontwikkeld door the Institute of Historical Research, University of London (in samenwerking met the British Library, the Oxford Internet Institute en Aarhus University). <https://www.webarchive.org.uk/shine> (17-04-2018). De ingevoerde zoekopdracht is in de onderstaande link te vinden: https://www.webarchive.org.uk/shine/graph?query=%22weblog%22%2C+%22blog%22&year_start=1996&year_end=2013&action=update (17-04-2018).

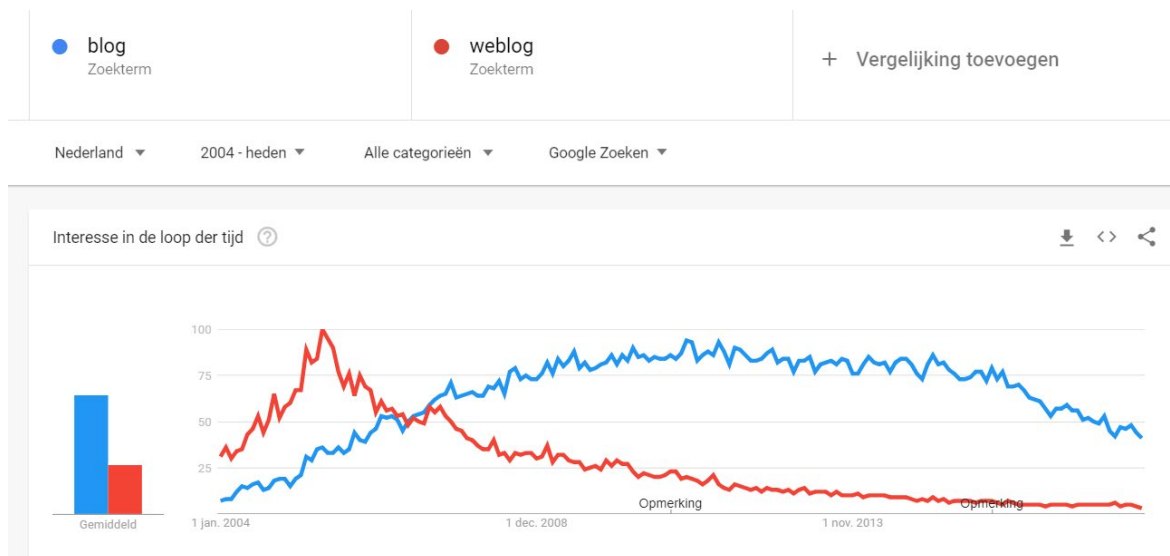
<https://web.archive.org/web/20180524123621/https://www.webarchive.org.uk/shine> (24-05-2018).

³⁸ <https://medium.com/google-news-lab/what-is-google-trends-data-and-what-does-it-mean-b48f07342ee8> (17-05-2018).

<https://web.archive.org/web/20180524130637/https://medium.com/google-news-lab/what-is-google-trends-data-and-what-does-it-mean-b48f07342ee8> (24-05-2018).

³⁹ https://trends.google.nl/trends/explore?date=all&geo=NL&q=blog_weblog (17-05-2018).

van de analyse geeft Google Trend ook aan dat de term 'blog' langzaamaan steeds minder vaak opgezocht wordt.



3 Digitaal erfgoed

3.1 Wat is de waarde van de collectie als digitaal erfgoed?

Het is van belang om weblogs te archiveren als digitaal erfgoed omdat het *born digital data* is. De webpagina's zijn niet in boeken te vinden en wanneer iemand zijn website verwijderd van de server is hij voor altijd weg. Het is uitzonderlijk als mensen hun website nog op opslagmedia bewaren (floppy, USB, CD, harde schijf) en nog moeilijker om een dergelijke versie te reconstrueren vanaf deze (soms verouderde) opslagmedia.⁴⁰ De gemiddelde internetpagina is bovendien slechts 90 dagen online: wat een kleine tijdspanne geeft voor het archiveren van de pagina.

Een veel gebruikt argument om internetwebsites en internetpagina's niet te archiveren is dat het Internet Archive dit al doet. Het Internet Archive (IA) archiveert inderdaad webpagina's, maar doet dit fragmentarisch. De gereconstrueerde websites zijn vaak een patchwork van stukjes en beetjes die op verschillende momenten zijn gearcheeerd. Zo wordt de hoofdpagina vaak regelmatig gearcheeerd, maar wanneer dieper in de website gezocht wordt, werken pagina's vaak niet of zijn onderdelen van de pagina's in een compleet ander jaar gearcheeerd. Dit geeft een verdraaid beeld van wat geesteswetenschappers in de toekomst zullen zien als een primaire bron. Als onderzoeker wil je dat de website met een vaste regelmaat en op één moment zo compleet mogelijk gearcheeerd wordt zodat de onderzoeker een goed beeld krijgt van hoe een website er op dat snapshot (één bepaald moment) heeft uitgezien. En laat dat nu precies zijn hoe de KB haar webarchief opbouwt.

Bij weblogs in het bijzonder is het van belang dat de Koninklijke Bibliotheek een zo compleet mogelijke archivering uitvoert omdat er veel archieven op de websites staan. Wanneer deze weblogs incompleet gearcheeerd worden, missen links naar eerdere posts waarop voortgeborduurd wordt. Een weblog is een ingewikkeld web van pagina's die naar elkaar verwijzen en in een unieke chronologische volgorde zijn opgebouwd. Zonder de juiste archiveringsprocedure kan er voor onderzoekers in de toekomst veel cruciale informatie en context verloren gaan.

Tenslotte gaat de wereld van de historische weblogs uit de jaren '90 snel ten onder. Er zijn bijvoorbeeld nog slechts twee van de 24 websites van grondleggers van de Nederlandse Blogosfeer online te vinden.⁴¹ Van de websites die nog online zijn, Profilic en L-rs, is alleen L-rs nog in zijn (min of meer) originele staat. Profilic maakt gebruik van de nieuwste WordPress toepassingen. Een platform dat tijdens de opkomst van de blogs nog niet gebruikt werd. Als historica gaan dan gelijk bij mij alle alarmbellen af. Om een vergelijking te maken: hoe beoordeel je een ouderwets middeleeuws manuscript dat eruit ziet als een modern geprint Harry Potter boek? Om in de toekomst goed onderzoek te kunnen doen, is het van belang dat we de veranderingen in vorm, software/programmeer-code en inhoud van het blog goed documenteren.

⁴⁰ Lees bijvoorbeeld dit stuk over het lastige reconstructieproces van het NL-menu. Kniiff, J. van der, 'It's Alive! Eerste Nederlandse webindex gered van CD-ROM', website *Netwerk digitaal erfgoed* (18-06-2018) <http://www.netwerkdigitaalerfgoed.nl/its-alive-eerste-nederlandse-webindex-gered-van-cd-rom/>

⁴¹ Meeuwssen, *Bloghelden*, 47. Geschreven rond 25-06-2018. Ten tijden van 2020 is er nog maar één grondlegger online; L-rs <http://l-rs.org/>, voor het laatst gewijzigd op 26-10-2010.

4 Onderzoeksmethode voor selectie

Aan het begin van het vormen van de collectie heb ik mij eerst ingelezen. De werken die vervolgens het meest invloed hebben gehad op het vormen van de collectieverantwoording, de definitie 'blog' en de indeling in thema's en format zijn:

- Bekker, C., *De blog bijbel. Voor beauty chicks & shopaholics* (Maartensdijk 2014).
- Deriemaeker, K., *Blog boek. Tips & inspiratie voor meer blog- en vlogplezier* (Antwerpen 2017).
- Duval, S., *How blogs work. Tips & tricks voor beginnende en ervaren bloggers* (Tielt 2014).
- Holthuijsen, D., *Weblogs. Uw persoonlijke blog op het web!* (Haarlem 2005).
- Meindersma, C., *Wetboek voor bloggers. De (on)geschreven regels voor bloggers, journalisten en content creators* (Den Haag 2016).
- <https://nl.wikipedia.org/wiki/Weblog> geraadpleegd op 9 april 2018.
- Meeuwssen, F., *Bloghelden* (2010).

Na het lezen van de literatuur heb ik de definitie bepaald van een 'blog' voor deze collectie. Deze definitie is gevormd naar aanleiding van de onderdelen van definities in onderzochte literatuur (meer hierover in paragraaf 5.2.). Aan de hand van deze definitie zijn de blogs geselecteerd en opgenomen in de collectie. Tijdens het vormen van de collectie is de definitie nog hier en daar aangepast vanwege hoofdzakelijk het probleem 'wanneer definieer je een post als nieuwsupdate en wanneer is het een blogpost?' (meer hierover in paragraaf 5.2.). In hoofdstuk 8 wordt uitgebreid beschreven welke informatie er in de collectie zit.

Anne Helmond & Esther Weltevrede

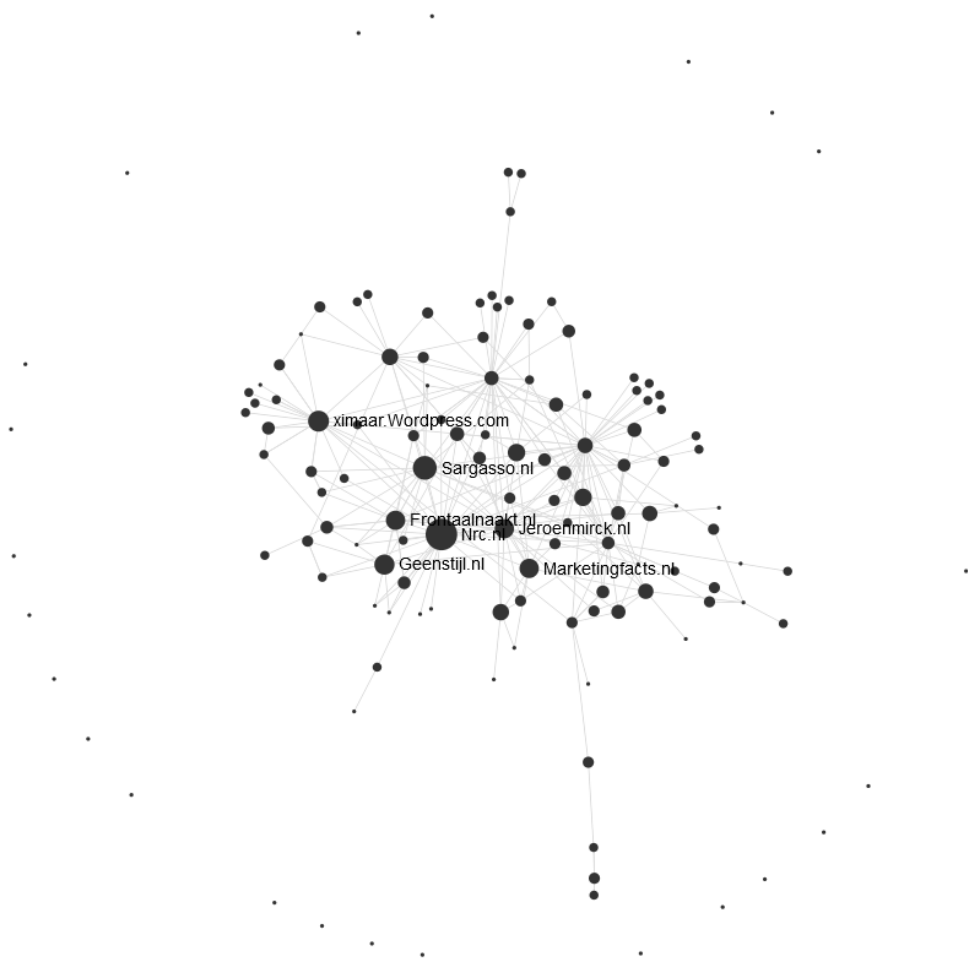
Oorspronkelijk was de bedoeling dat ik 'kritisch zou gaan kijken naar de selectie en reconstructie van Helmond en Weltevrede en vervolgens de collectie zou uitbreiden, analyseren, beschrijven en archiveren.'⁴² Omdat er in de eerste helft van mijn stage wat verwarring was over de Helmond lijst (ik had een opgeschoonde literatuurlijst in plaats van de dataset), ben ik hier pas halverwege mijn stage mee aan de slag gegaan. Het gevolg hiervan is dat de NL-blogosfeer collectie andere blogs bevat dan de Helmond lijst, waardoor de twee lijsten elkaar aanvullen. Beiden onderzoeken kwamen echter wel uit bij *Bloghelden* als belangrijke bron voor een onderzoek naar de blogosfeer. Hierdoor zit er ook overlap in de onderzoeken.

Mijn analyse van de Helmond lijst is uiteindelijk het volgende: de lijst moest ontduddeld worden voor ik ermee aan de slag kon. Van de 2500 records heb ik het nu teruggebracht naar de 1500. De reden dat er dubbele records waren, was waarschijnlijk doordat er een spatie of slash (/) teken aan het eind van sommige internet adressen stond en het verschil in [http://www.](http://www) en alleen <http://> aan het begin van sommige hyperlinks. Deze lijst zal een 'bonuslijst' worden omdat er geen tijd is deze 1500 websites kwalitatief te beoordelen zoals met de huidige collectie wel is gedaan. De Helmond lijst geeft een

⁴² Gebaseerd op het onderzoek: Weltevrede, E. and Helmond A. (2012) Where do bloggers blog? Platform transitions within the historical Dutch blogosphere. *First Monday* 17(2). DOI: [10.5210/fm.v17i2.3775](https://doi.org/10.5210/fm.v17i2.3775).

interessant beeld van de Nederlandse blogosfeer ondanks dat bijvoorbeeld de Blogparel wedstrijden hier niet genoemd worden. In de Helmond lijst zijn ook Belgische blogs opgenomen omdat experts (bloggers die zij interviewden) meenden dat zij deel uitmaakten van de Nederlandse blogosfeer.

Op 15 mei 2018 had ik nog een afspraak met Anne Helmond over haar onderzoek. De onderzoekslijst van blogs die Helmond & Weltevrede hebben gebruikt in hun onderzoek is gebaseerd op de mening van de expert bloggers in Nederland (de bloghelden/ Dutch bloggies kring).⁴³ Het onderzoek van Helmond & Weltevrede vond plaats rond 2011 waardoor zij weinig last hebben gehad van het Sanoma incident. Toen ik de Blogparel wedstrijd noemde, had Anne nog niet van deze wedstrijd gehoord. Het gevaar met de expertkring is dat je groepen mist. Na de vergadering met Anne heb ik een linkanalyse uitgevoerd om te kijken in hoeverre de Blogparel kring en de Bloghelden kring verbonden waren met elkaar. Uit de linkanalyse met Hyphe bleek dat de websites (van de prijswinnaars/genomineerden) van de twee wedstrijden redelijk met elkaar verbonden zijn. Er wordt veel naar elkaar verwezen waardoor er geen twee aparte sferen te zien zijn.



Linkanalyse Hyphe: Dutch Bloggies en Blogparels

⁴³ Weltevrede, E. and Helmond A. (2012) Where do bloggers blog? Platform transitions within the historical Dutch blogosphere. *First Monday* 17(2). DOI: [10.5210/fm.v17i2.3775](https://doi.org/10.5210/fm.v17i2.3775).

Anne is als onderzoeker geïnteresseerd in hoeveel versies de KB heeft van een gearcheverde website. En daarnaast hoeveel pagina's er dan (per versie) gearcheverd zijn. Als er een archief aanwezig is, kan dit bijvoorbeeld ook een goed beeld geven van het aantal posts. Daarnaast vroeg ze meteen wat de archiveerfrequentie was (hoe vaak per jaar een website gearcheverd werd). Dit is belangrijk om goed aan te geven in de collectiebeschrijving. Qua terminologie van de kolommen in de NL-blogosfeer dataset dacht Anne dat webadres / URL beter het webdomein kon heten. De webpagina is dan de URL. Er is uiteindelijk besloten de huidige terminologie aan te houden.

Anne Helmond was ook aanwezig tijdens een korte lunchpresentatie die ik hield op 27 juni 2018.

5 Selectiemethode

5.1 Selectiecriteria voor archivering (Koninklijke Bibliotheek)

De KB heeft als selectiebeleid het verzamelen van websites op het gebied van Nederlandse taal, geschiedenis en cultuur.⁴⁴ Deze collectie bevat daarom uitsluitend Nederlandse blogs. Wij verstaan onder Nederlandse blogs:

- Blogs die in het Nederlands zijn geschreven
- Blogs met Nederland als onderwerp
- Blogs geschreven door een Nederlander
- Blogs geschreven door een Nederlander die momenteel in een ander land is. (reisblogs, expats)

Wanneer een Nederlander op een buitenlandse website blogt, moet dit wel een significante bijdrage zijn om opgenomen te worden in de collectie. Enkele posts op een buitenlandse website zijn niet voldoende om opgenomen te worden in de Nederlandse blogosfeer.

- Geschreven door een buitenlander (die in Nederland woont) zoals expats.
- Domeinnamen (blogs van het Nederlandse webdomein)

De KB heeft ontzettend veel websites met het Nederlandse domein: de .nl websites. Andere domeinen die worden gearchiveerd zijn: .com, .net, .eu, .nu en .org. Buitenlandse domeinnamen zijn in de NL-blogosfeer collectie van de KB wel ondervetegenwoordigd.

De NL-blogosfeer collectie moet een representatief beeld geven van de Nederlandse samenleving en de digitale cultuur op het Nederlandse deel van het wereldwijde web. Daarnaast worden websites opgenomen vanwege hun nationaal cultureel historisch belang.

Wat wordt niet gearchiveerd door de KB?

- Websites met digitale spelletjes
- Online nieuws wordt vanaf 2018 gearchiveerd met uitzondering van:
 - GeenStijl wordt gearchiveerd vanaf 8 februari 2013
 - Joop wordt gearchiveerd vanaf 7 januari 2014
 - Nu.nl wordt gearchiveerd vanaf 7 september 2011 dagelijks de startpagina en één vervolgpagina (op het middaguur). Vanaf 3 maart 2015 wordt de website daarbovenop jaarlijks volledig gearchiveerd.
- Webcams en andere live content

⁴⁴ 'Selectie bij webarchivering', Website Koninklijke Bibliotheek | Nationale Bibliotheek van Nederland (26-11-2020)
<https://www.kb.nl/organisatie/onderzoek-expertise/e-depot-duurzame-opslag/webarchivering/selectie-bij-webarchivering>

- De uitzondering hierop zijn de liveblogs op nu.nl. Deze worden gearchiveerd mits het liveblog op het middaguur online is en op de startpagina of één niveau diep staat. Als de liveblogs dieper in de website zitten, dan worden ze jaarlijks gearchiveerd.
- MSN
- Datasets (dit doet DANS)⁴⁵
- Bulletin boards, prikbordsites met bijdragen van derden
- Intranets
- Radio- en televisieprogrammasites (websites van omroepen): Beeld & Geluid archiveert deze websites.
- Sociale media websites (Twitter, Facebook, LinkedIn, Tumblr, Instagram, Snapchat, Pinterest, etc.)

Social media is een belangrijk medium waarop met name microblogs verspreid worden. Het is echter voor de KB op het moment juridisch en technisch niet mogelijk om deze websites (of apps zoals Instagram en Snapchat) te archiveren. Een Youtube Vlogster zoals Famke Louise was hierdoor wel geselecteerd voor de collectie maar haar kanaal kan niet gearchiveerd worden.

- Commerciële websites (tenzij er een gegronde reden voor is)
- Websites van politieke partijen en politici (Archipol in Groningen archiveert deze websites)

Archipol archiveert alleen websites van politieke partijen en politici. In het algemeen politiek gekleurde websites archiveert Archipol dus niet. Deze archiveert de KB wel. Wanneer er blogs van politici of politieke partijen gevonden zijn, wordt aan Archipol gevraagd of zij deze willen archiveren als de website nog bestaat. Dit is tot nu toe nog niet nodig geweest omdat zij de blogs van politici zoals <http://www.janmarijnissen.nl/index.html> en <http://www.gerritzalm.nl/> al in hun database hebben zitten. De website die ze niet hebben, is <http://www.remcokouwenhoven.nl/> maar omdat deze website niet langer online is, kan deze niet gearchiveerd worden.

Risicovolle websites (abjecte websites):

Websites met een politiek of religieus extreem karakter, onacceptabele of zelfs illegale inhoud worden niet opgenomen in de selectie. Dit zijn bijvoorbeeld websites van extreme politieke groeperingen, criminele organisaties en maatschappelijk onacceptabele verenigingen. De KB heeft deze keuze gemaakt omdat het erkennen van deze organisaties of personen wil voorkomen, het de wet niet wil overtreden of om zichzelf en anderen in bescherming te nemen. Daar komt nog bij dat vaak de rechthebbende persoon van een dergelijke website niet te achterhalen is omdat zij onder een pseudoniem publiceren.

⁴⁵ Data Archiving and Networked Services, <https://dans.knaw.nl/nl> (28-10-2020).

5.2 Selectiecriteria voor NL-Blogosfeer collectie

Definitie 'blog'

Mijn definitie van een blog is gevormd naar aanleiding van de onderdelen van definities in onderzochte literatuur. Kijk voor de elementen en literatuur die gebruikt zijn in bijlage 2. Per boek staat aangegeven welk element gebruikt is of dat er een soortgelijke term gehanteerd is.

Een blog is:

- Een reeks afzonderlijke berichten (posts)
 - Eén blogpost is ook mogelijk (maar uitzonderlijk)
- Waarbij een omgekeerde chronologische volgorde wordt aangehouden: het meest recente bericht staat bovenaan.
- Er verschijnen regelmatig nieuwe berichten: het liefste dagelijks of wekelijks.
- Een blog, of de officiële term: een **weblog**, staat online op een aparte, eigen, website of is een onderdeel van een website.
 - De website kan een eigen domeinnaam zijn of onderdeel van een weblog service zoals WordPress, BlogOnline of Blogspot.
- Bevat met name tekst, maar kan ook alleen links, afbeeldingen (photolog), video (vlog) of geluid (podcast) zijn.
- Een blog kan ook een combinatie zijn van tekst, afbeeldingen, geluid, (interne en externe) links en video's.
- De doelgroep van een blog staat niet vast / kan veranderen.
- Blogs kunnen openbaar of gesloten zijn. Gesloten blogs zijn bijvoorbeeld een blog binnen intranet of een blog die alleen op uitnodiging te lezen is (<http://estherdonkers.nl/blog/> oudere posts onderaan) of een blog wat achter een betaalmuur zit (zoals <https://boekblad.nl/Blogs/>).
- Iedereen kan bloggen: van amateurs tot professionals. Je kunt er zelfs je beroep van maken.
- Een blog kan door één of meerdere personen geschreven worden. Vaak staat bij een blog dat door meerdere auteurs wordt geschreven de auteur per post genoemd.
- Er is iemand (of meerdere personen) verantwoordelijk voor de inhoud. Niet iedereen kan zomaar berichten (posts) plaatsen op een blog.

Veel voorkomende elementen (maar geen verplichte onderdelen)

- Bezoekers kunnen meestal reageren op een blog
- Blogs zijn meestal onderling met elkaar verbonden
- Een blog kan een gids zijn die de lezer mee op sleeptouw neemt om de lezer dingen te tonen die hij voor jou heeft uitgekozen.⁴⁶
- Een blog kan een archief hebben waarin alle bijdragen bewaard worden

⁴⁶ Deriemaeker, K., *Blog boek. Tips & inspiratie voor meer blog- en vlogplezier* (Antwerpen 2017) 24.

- Tijd en datum staan meestal bij de post genoemd

Is het nieuws of is het een blog?

Eén van de vraagstukken waarmee wij worstelden was: wanneer is iets een blog en wanneer is iets een nieuwsrubriek? Een nieuwsrubriek is namelijk niet hetzelfde als een blog. Wanneer een website een twijfelgeval was, dus of het een nieuwssite/tijdschrift of een blog was, heeft de reactiemogelijkheid **op de website zelf** (dus geen reacties via bijvoorbeeld Facebook) doorslag gegeven.

Blogplatforms

Overkoepelende blogplatformen hebben we niet als geheel meegenomen. Het kan wel zo zijn dat individuele blogs zijn aangeschreven (die onderdeel zijn van zo een platform). Voorbeelden van blogplatformen zijn:

<http://blogonline.blog.nl/over-blognl> ⁴⁷

<http://www.alleweblogs.nl/> ⁴⁸

<http://blog.nl/wat-blog-nl> ⁴⁹

<https://tweakblogs.net/>

5.3 Excel uitleg: bronbestanden / geselecteerd via

Uitleg over de verschillende bronbestanden en de selectiemethode genoemd in het Excelbestand NL-blogsfeer in de kolommen 'bronbestand' en 'geselecteerd via'.

Geselecteerd via:

Handmatig

Deze records zijn toevallig geselecteerd omdat iemand hem vond en het een interessante toevoeging leek aan de collectie. Er staat dan vaak een X bij bronbestand omdat hij rechtstreeks in het eindbestand is gezet. Dit is echter niet altijd het geval. Wanneer er handmatig staat met een bronbestand zijn de links handmatig geselecteerd na beoordeling van de website tijdens het werken / beoordelen van websites uit dat bronbestand.

Zoekterm X Deze websites zijn gevonden door het gebruik van een zoekquery.

Dutch bloggie winnaar Winnaar van een Dutch Bloggie award.

⁴⁷ <https://web.archive.org/web/20180531074216/http://blogonline.blog.nl/over-blognl> (31-05-2018).

⁴⁸ <https://web.archive.org/web/20180531074431/https://www.alleweblogs.nl/> (31-05-2018).

⁴⁹ <https://web.archive.org/web/20180611092240/http://blog.nl/wat-blog-nl> (11-06-2018).

Literatuur Website is geselecteerd naar aanleiding van literatuur onderzoek
NL menu (2004) weblog tag Zie: NL menu onder 'bronbestanden', volgende pagina.
Blogparel jaartal X Genomineerde bij de Blogparel wedstrijd.
Blogger des Vaderlands X Jurylid, genomineerde of winnaar van de Blogger des Vaderlands wedstrijd. Ook de pagina van de wedstrijd zelf is een blog.
TPO artikel Blogs geselecteerd naar aanleiding van het lezen van het volgende artikel⁵⁰: <http://tpo.nl/2010/12/28/verplicht-in-uw-rss-reader-Weblogs-die-er-echt-toe-doen/>

Ongeldige links (al in KB webarchief)

Links die niet werkten of twijfelachtig waren zijn aan de start van de collectievorming apart gehouden. Er is vervolgens gecontroleerd of ze al in het KB webarchief zaten (en beoordeeld konden worden) en of het zeker weten blogs waren. Deze zijn alsnog toegevoegd met de voorgaande tag omdat zij uit het bestand 'ongeldige links' komen.⁵¹

Blogawards HU Blogs geselecteerd naar aanleiding van de blogawards van (& voor) studenten van de School voor Journalistiek. <http://www.journalismlab.nl/tag/blogawards/> De blogs zijn geschreven door studenten voor hun journalistiek opleiding. De beste blogs werden genomineerd voor een prijs. Wanneer deze blog na het voltooien van het vak is gestopt, is het blog ingedeeld onder het thema educatie omdat zij het blog schreven voor hun opleiding. Is het blog na het voltooien van het vak nog wel bijgehouden dan kan het een ander thema krijgen.

Bronbestanden:

X Dit record heeft geen bronbestand maar is rechtstreeks in het eindbestand ingevoegd.

Helmond links Dit zijn links die uit de literatuurlijst van de paper van Helmond komen.⁵²

Helmond dataset Alleen van toepassing in de bonuslijst Helmond Dutch blogosphere schoon en ontdubbelt (zie Hoofdstuk 10 Datasets).⁵³

Vinexlijst

Elke maand publiceren de Verenigde Internet Exploitanten (VINEX) de top 50 mediamerken op basis van bereik (6+). Bovendien wordt dit bereik uitgesplitst naar de devices desktop/laptop, smartphone

⁵⁰ <https://web.archive.org/web/20180528082618/http://tpo.nl/2010/12/28/verplicht-in-uw-rss-reader-Weblogs-die-er-echt-toe-doen/> (28-05-2018).

⁵¹ Meer uitleg over deze bonuslijst staat in Hoofdstuk 10 Datasets.

⁵² Weltevrede, E. and Helmond A. (2012) Where do bloggers blog? Platform transitions within the historical Dutch blogosphere. *First Monday* 17(2). DOI: [10.5210/fm.v17i2.3775](https://doi.org/10.5210/fm.v17i2.3775).

⁵³ Weltevrede, E. and Helmond A. (2012) Where do bloggers blog? Platform transitions within the historical Dutch blogosphere. *First Monday* 17(2). DOI: [10.5210/fm.v17i2.3775](https://doi.org/10.5210/fm.v17i2.3775).

en tablet.⁵⁴ Niet de hele lijst is opgenomen. Er is per website beoordeeld of er een blog element op te vinden was. Het is op dit moment nog onduidelijk welk VINEX bestand hiervoor gebruikt is. Het is waarschijnlijk geen top 50 bestand (het bestand wat ik heb gebruikt heeft meer dan 400 records). Eerder een mediamerkenlijst. De datum van uitgave van de lijst is ook nog onbekend.

Dutch Bloggies Prijs

Informatie over de winnaars van de Dutch bloggies komt uit het boek *Bloghelden* en de Wikipedia pagina. De winnaars kunnen meerdere bloggies hebben gewonnen.

Links uit literatuur

Deze blogs zijn gevonden tijdens het lezen van literatuur over blogs. Ze werden daarin als voorbeeld genoemd.

NL menu

De Koninklijke Bibliotheek heeft het NL menu (versie uit 2004) gereconstrueerd van een CD af.⁵⁵ Na zoeken op deze gereconstrueerde website vond ik het weblog van Remco Kouwenhoven. Na het drukken op het vierkantje met de i erin vond ik de rubriek weblogs. Door op deze tag te drukken kreeg ik een lijst met weblog getagde websites. Deze heb ik in een Excelbestand gezet en vervolgens in de collectie opgenomen.

Blogger des Vader. (etc):

1. Blogger des Vaderlands

In 2015 werd een wedstrijd gehouden om een Blogger des Vaderlands te benoemen. Dit was een initiatief van het Blogbal.⁵⁶

2. Blogparel <http://blogparel.nl/>

De Nederlandse Nobele Prijs voor Blogs, georganiseerd door Peter de Jonge (blogger alias: P. J. Cokema).

⁵⁴ <https://vinex.nl/resultaten/> <https://web.archive.org/web/20180528082940/https://vinex.nl/resultaten/> (28-05-2018).

⁵⁵ Kniiff, J. van der, 'It's Alive! Eerste Nederlandse webindex gered van CD-ROM', website *Netwerk digitaal erfgoed* (18-06-2018) <http://www.netwerkdigitaal.erfgoed.nl/its-alive-eerste-nederlandse-webindex-gered-van-cd-rom/>

⁵⁶ <http://hetblogbal.blogspot.nl/p/blogger-des-vaderlands.html> Opgenomen in het webarchiveringsproces van de KB op 04-06-2018; eerste archivering begin juli. <https://web.archive.org/web/20180528083552/http://hetblogbal.blogspot.nl/p/blogger-des-vaderlands.html> (28-05-2018).

- De pagina van 2009 was niet langer meer online en is daarom (tot nu toe) niet meegenomen.
- De Blogparel wedstrijden van 2010⁵⁷, 2011⁵⁸ en 2013⁵⁹ zijn wel meegenomen in de collectie.
- De websites (blogs) van de wedstrijd zelf zijn ook gearcheveerd.

KB blogs Blogs gevonden in de KB webcollectie. (Vaak niet traceerbaar in link).

Kleine lijstjes TPO artikel & Mom blog awards⁶⁰

⁵⁷ <http://www.peterspagina.nl/blogparel/>

Gearchiveerd door de KB op 09-05-2018

⁵⁸ http://www.peterspagina.nl/blogparel_2011/

Gearchiveerd door de KB op 09-05-2017

⁵⁹ <http://www.blogparel.nl/2013/index.php/de-jury-aan-de-slag/>

Gearchiveerd door de KB op 09-05-2017

⁶⁰ <https://web.archive.org/web/20180528082618/http://tpo.nl/2010/12/28/verplicht-in-uw-rss-reader-Weblogs-die-er-echt-toe-doen/> (28-05-2018). <https://web.archive.org/web/20180412130350/http://www.momblogawards.nl/> (12-04-2018).

6 Contactadres, analoge contacten

6.1 Contactadres / verweesde websites

Vanwege het format blog zijn er een relatief groot aantal websites zonder contactadres. Dit kan bijvoorbeeld komen omdat een blog geschreven is onder een pseudoniem. De eigenaren kunnen door een ontbrekend contactadres niet benaderd worden met een opt-out brief en deze websites vallen buiten archivering.⁶¹

31 mei 2018: 22 websites hebben tot nu toe geen contactadres.

20 juni 2018: 42 websites hebben tot nu toe geen contactadres.

⁶¹ Dit worden ook wel verweesde websites genoemd: websites waarvan de rechthebbende niet te achterhalen is.

7 Techniek en technische problemen

7.1 WCT / Heritrix

Gebruikte software

2007 – 2009: versie 1.10.1. van Heritrix gebruikt, aangestuurd vanuit Linux.

Vanaf 2009: versie 1.14.1 van Heritrix gebruikt in combinatie met de administratieve tool: WebCuratorTool versie 1.4. De websites die voor 2009 zijn gearchiveerd zijn wel gearchiveerd en worden getoond in een aangepaste versie van de WayBack Machine maar missen informatie over de provenance en de authenticiteit door het gebrek aan metadata.

De KB gebruikt een aangepaste versie van de WayBack Machine voor de beschikbaarstelling voor het publiek.

Archiveringsproces

Na selectie & keuring door mij worden de websites in de WebCuratorTool (WCT) ingevoerd. Dan wordt er vervolgens een opt-out verstuurd naar de eigenaar van de website. De eigenaar heeft vervolgens een maand om te weigeren. Weigert hij of zij niet dan wordt de website voor de eerste keer gearchiveerd. De meeste websites worden vervolgens jaarlijks gearchiveerd (met als enkele uitzondering nu.nl, die dagelijks gearchiveerd wordt of evenementenwebsites van bijvoorbeeld het sleepwet referendum). Voor blogs geldt dit ook: zij worden over het algemeen jaarlijks gearchiveerd. Na ongeveer een jaar kunnen de archiveerde websites ingezien worden in de (aangepaste) WayBack Machine van de KB op computers in de leeszaal van de Koninklijke Bibliotheek. Hiervoor is een jaarpas abonnement van de KB nodig.

7.2 Technische problemen

Sommige websites (weblogs) zijn niet te archiveren vanwege technische problemen.

Javascript: Het is niet mogelijk om complexe Javascript (Javascriptjes) te archiveren. Deze worden daarom niet gearchiveerd.

Externe verwijzingen naar social media: Vlogs (videoblogs) en podcasts (audio blog) kunnen alleen gearchiveerd worden wanneer deze binnen het domein van de website staan. Wanneer naar een externe website als YouTube verwezen wordt, kan dit element niet gearchiveerd worden. Dit geldt ook voor posts die verwijzen naar social mediasites zoals lijsten met tweets. Ook deze onderdelen kunnen niet gearchiveerd worden.

Dark web: Niet aanwezig. Deze websites hebben zichzelf onvindbaar gemaakt en zijn niet toegankelijk voor standaard browsers of webarchiveringtools.

7.3 Gebruikersproblemen

Problemen voor onderzoekers in de collectie

Contextverlies

Blogwebsites worden in de KB meestal jaarlijks gearchiveerd. Hoewel dan wel het hele archief wordt gearchiveerd met alle blogposts, mist de onderzoeker de oorspronkelijke context waarin de posts zijn gepubliceerd. Bijvoorbeeld: het is niet meer te achterhalen welke andere artikelen (of blogposts) rondom de post stonden. Daarnaast kunnen eventuele afbeeldingen/video's (etc.) missen omdat het blog in het archief in kleiner formaat wordt opgeslagen (gecomprimeerd bijgevoegd worden) dan het origineel. Dit is vaak om bijvoorbeeld archiefruimte te besparen. Dagelijks archiveren van websites is technisch niet haalbaar. Een voorbeeld van missende afbeeldingen is te zien in de Wayback Machine van het Internet Archive via de website Rumag.⁶²

Verkochte webdomeinen

Niet alle webadressen gaan offline wanneer een blog stopt. Sommige webadressen worden doorverkocht. Sommige van deze websites werken dan nog wel maar zijn heel andere websites. Zo is het Lines & Splines blog nu een Duitse casino site.⁶³ Bij een linkanalyse (van het live web) kan hierdoor een probleem ontstaan in de analyse omdat een geldige link wordt verwerkt maar de verkeerde website. Zo verwijzen tegenwoordig heel veel x.web-log.nl of weblog.nl domeinen door naar <https://www.sanoma.nl/>. In de kolom 'website verkocht / link verwijderd' is onder andere in 2018 genoteerd welke websites links zijn overgenomen door een andere website. Daarna is dit echter niet meer bijgewerkt.

Juridische beperkingen

Website eigenaren hebben het recht om archivering van hun weblog te weigeren. Hierdoor kan het zijn dat een blog niet gearchiveerd is. Zij kunnen weigeren omdat Nederland (& Zwitserland als enige twee landen in de wereld) geen depotverplichting heeft. In Nederland geldt als regeling een nationaal depot waarbij op vrijwillige basis wordt gewerkt.

⁶² <https://web.archive.org/web/20180619112324/https://rumag.nl/met-deze-emojis-fix-jij-elke-jongen-of-meisje-via-whatsapp/> (19-06-2018).

⁶³ Hofstede, M., *Het zoekboek voor het web* (Rijswijk 2001) 58. **Pas op wanneer je op deze link klikt! De site is immers overgenomen en nu een casino site.** <http://www.linesandsplines.com/> (01-05 -2018).

Social media (websites/apps) kunnen niet gearchiveerd worden vanwege juridische (en technische) beperkingen. Nederland heeft zoals hierboven al genoemd namelijk geen wettelijk depot. Van elke website die geselecteerd wordt, moet de KB de eigenaar benaderen met een opt-out.

Tenslotte is het daadwerkelijk inzien van het webarchief beperkt voor gebruikers. Wanneer u het webarchief wilt inzien moet u naar de KB komen om in de leeszaal van de KB de websites te bekijken in de Wayback Machine. U heeft hiervoor een KB-jaarpas nodig. Voor meer informatie over de gebruiksvoorwaarden van het webarchief van de KB (ten tijde van 28-05-2018) zie de volgende voetnoot.⁶⁴

⁶⁴ <https://www.kb.nl/bronnen-zoekwijzers/databanken-mede-gemaakt-door-de-kb/webarchief-kb/gebruiksvoorwaarden-webarchief-koninklijke-bibliotheek> (28-05-2018). <https://web.archive.org/web/20180528133932/https://www.kb.nl/bronnen-zoekwijzers/databanken-mede-gemaakt-door-de-kb/webarchief-kb/gebruiksvoorwaarden-webarchief-koninklijke-bibliotheek>

8 Geselecteerde sites en beschrijving topstukken

8.1 Kolommen

De NL Blogosfeer collectielijst met handmatige metadata is een lijst van blogs in Excel met informatie over het blog. Hieronder volgt een overzicht van de kolommen (informatie) die in de collectie zijn opgenomen. De **kolomnamen** staan dikgedrukt en zijn toegelicht waar nodig over wat er precies in de kolommen staat.

Kolommen die informatie bieden over het archiveerproces

- **in WCT gezet (datum)** WCT = WebCuratorTool
- **al in WCT (datum)**

De KB geeft bij blogs geen einddatum op bij de selectie. Wanneer een website geselecteerd wordt, wordt hij in principe gearchiveerd tot de website niet meer live is.

Kolommen die informatie bieden over het selectieproces

- **Geselecteerd via**
- **Bronbestand** (voor meer informatie zie: paragraaf 5.3)

In deze kolommen is gezet met welke zoekterm of welke methode de website is geselecteerd. Dit zijn twee kolommen waaruit duidelijk wordt uit welk bestand de link komt (literatuur of de uitdraai van de webcollectie) en hoe het blog gevonden is, bijvoorbeeld door de zoekterm 'blog' of de zoekterm 'wordpress'. Daarnaast is de term 'handmatig toegevoegd' gehanteerd voor websites die wij toevallig tegen het lijf liepen. Deze kolommen helpen de gebruiker begrijpen waarom en hoe wij bepaalde websites hebben geselecteerd.

Kolommen die informatie geven over de website zelf

- **Webadres (URL)**

De KB archiveert de hele website, zelfs wanneer er slechts één pagina van de website een blog is. De URL wordt daarom altijd opgenomen in het Excelbestand. Wanneer een blog slechts een deel van de website is, wordt de URI nog eens apart genoemd en opgenomen in de kolom 'webpagina (URI)'.

Op dit moment (19 juni 2018) is er voor één website alleen een URI met het blog opgenomen:

<https://blog.huislijn.nl/>

Het deels archiveren van een website leidt wel tot contextverlies (namelijk de rest van de website).

- **Webpagina (URI)**

De collectie bevat in principe hele websites. Er wordt echter een tweede kolom opgenomen voor wanneer een blog zich als apart onderdeel binnen een website bevindt. Wanneer de hele website een blog is, blijft deze kolom leeg.

- **Webadres verkocht / link verwijderd**

Oude website URL's zijn soms verkocht aan andere personen waardoor de link nog wel bestaat, maar het niet langer de website is die wij bedoelen. Daarnaast kunnen websites helemaal opgeheven zijn. Er komt een X te staan wanneer dit het geval is. **Let op** dat deze kolom sinds 2018 niet meer is bijgewerkt.

- **Speciale collectie**

Tag toevoegen van deze collectie: NL-blogsfeer.

- **Andere speciale collecties**

Een tag naar andere relevante speciale collecties van de KB waartoe de website behoort. Hier kan de gebruiker extra informatie vinden over het blog.

- **Naam** Naam van de website

Idealiter overgenomen uit de maandelijkse uitdraai van alle websites die in de WCT van de KB staan. Op deze manier zijn collecties in de toekomst makkelijk samen te voegen via bijvoorbeeld MySQL Workbench. Niet overtypen, letterlijk de cel kopiëren. De kleinste typfouten (zoals spaties) zijn al catastrofaal. Echter: vele websites zaten nog niet in de KB database. We doen ons best de gegevens in deze collectie (kolom) zo identiek mogelijk te houden met de naam kolom in de KB database (van alle gearchiveerde websites).

- **Thema / genre**

Dit is een dappere poging van mij om de blogs in te delen aan de hand van een hoofdthema. Voor meer informatie (en de problemen met deze kolom) zie paragraaf 8.3. Genre.

- **Datum blog gestart**

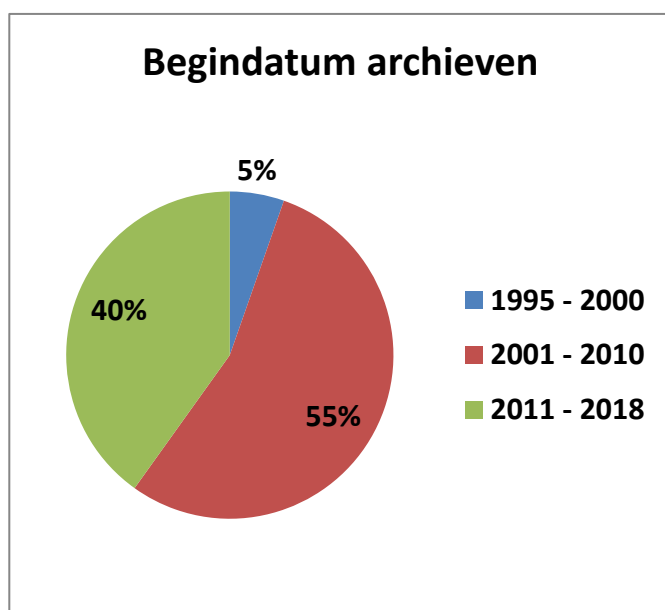
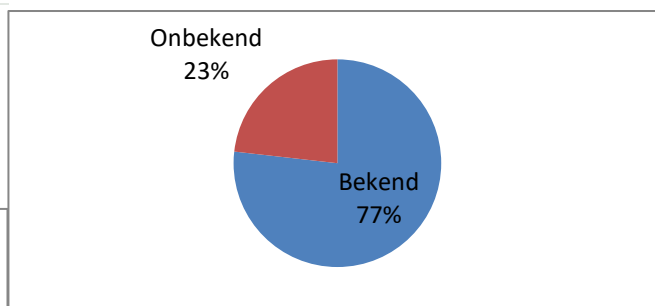
Dit is gedaan aan de hand van de volgende stappen:

- Eerst is gekeken naar de oudste datum van het archief
- Wanneer er andere informatie beschikbaar was over de oorsprongsdatum wordt de oudste datum in het archief vervangen
- Vanwege tijd is ervoor gekozen om wanneer er te lang teruggescrold moest worden in het archief, deze informatie niet op te zoeken maar 'onbekend' in te vullen.
- Eventueel kan bij deze 'onbekende' records nog gekeken op WHOIS SIDN.⁶⁵ Met dit webtooltje kan gekeken worden op welke datum een website is geregistreerd. Dit is echter nog niet gedaan.

In de kolom staat een jaartal gevolgd door maand (wanneer deze gegevens beschikbaar zijn).

Op 20-06-2018 waren dit de volgende concrete aantallen (waarbij in de analyse alleen het jaartal is meegenomen waarin de website is begonnen)⁶⁶:

Onbekend	107
1995 - 2000	19
2001 - 2010	193
2011 - 2018	142
Eindtotaal	461



⁶⁵ <https://www.sidn.nl/whois> (28-05-2018).

⁶⁶ Analyses in deze collectiebeschrijving zijn gemaakt op basis van de selectielijst van 20-06-2018. Deze selectielijst wijkt af van de huidige NL-blogsfeer collectielijst.

- **Meerdere auteurs**

Wordt het blog door één (hoofd)persoon (NEE) geschreven of door meerdere auteurs (JA)? Of is dit onduidelijk (onbekend).

- **Format**

Voor uitgebreidere informatie zie: paragraaf 8.2.

Indeling is als volgt:

- Tekst Tekstuele blog
- Vlog
- Fotoblog Photolog, phlog
- Podcast
- Linkdump Linklog
- Microblog Facelog / space blog
- MIX
- Moblog
- Onbekend

Overig

- **Index nr.**

Als het nodig is om de collectie in één bepaalde volgorde te houden is het belangrijk om een kolom met indexnummers te maken in kolom één. Daardoor kan je altijd weer terug naar de originele volgorde wanneer je het bestand even anders gesorteerd hebt omdat je op zoek was naar bijvoorbeeld alle tekstuele blogs of alle Nederlandse domeinen. Op dit moment staan de indexnummers op volgorde waarop de websites in het bestand zijn gezet. De NL Blogosfeer collectielijst met handmatige metadata staat standaard gesorteerd op index nr.

Kolommen die niet opgenomen zijn (en waarom)

- **Politieke kleur**

Deze categorie is niet opgenomen omdat dit voor blogs heel ingewikkeld is om te benoemen, zeker bij meerdere auteurs.

- **UNESCO code**

Deze code wordt al aangewezen wanneer een collectiespecialist de website invoert in WCT. Voor websites die de KB gearchiveerd heeft, is het daarom makkelijk om deze informatie over te nemen als een tweede thematische indeling, indien nodig.

Interessante kolommen maar niet mogelijk in deze tijdsspanne

- Indeling naar hoeveel mensen er precies aan het blog schrijven: specifieke aantallen.
- Taal In welke taal is het blog geschreven
- Aantal posts Dit is problematisch omdat live websites steeds veranderen. Het is mogelijk dit toe te voegen aan websites die niet langer live zijn. Bij deze websites zal dat aantal immers niet meer veranderen.
- Frequentie Hoe vaak wordt een website gearchiveerd door de KB. De standaard is jaarlijks voor blogwebsites.

De meeste blogs worden jaarlijks gearchiveerd. Wanneer we echter nieuwsblogs als nu.nl (wat dagelijks twee niveaus gearchiveerd wordt) opnemen in de collectie is het toch handig om te benoemen dat er ook andere frequenties zijn. Bovendien wordt soms voor de actualiteit een website vaker gearchiveerd.

8.2 Formats & indeling personen

Indeling naar gebruikers

- Individuele blog: één vaste auteur schrijft het blog.⁶⁷
- Groepsblog (Multi Author Blogs): meerdere vaste auteurs schrijven het blog.⁶⁸
- Gastblogs: blogs die door iemand anders geschreven zijn dan door de persoon of personen die normaal de content verzorgen.⁶⁹ Eventueel: blogs die continue van gebruiker wisselen zoals bij de WeAreXpats twitter.⁷⁰

Omdat deze indeling te tijdrovend was, is er gekozen voor de simpelere aanduiding: wordt dit blog door meerdere personen geschreven? Ja of Nee. Wanneer een individueel blog ook regelmatig gastposts bevat, valt dit ook onder een 'ja'; dit blog wordt door meerdere personen geschreven.

⁶⁷ Meindersma, C., *Wetboek voor bloggers. De (on)geschreven regels voor bloggers, journalisten en content creators* (Den Haag 2016) 33.

⁶⁸ Meindersma, C., *Wetboek voor bloggers. De (on)geschreven regels voor bloggers, journalisten en content creators* (Den Haag 2016) 33.

⁶⁹ Meindersma, C., *Wetboek voor bloggers. De (on)geschreven regels voor bloggers, journalisten en content creators* (Den Haag 2016) 34.

⁷⁰ <http://wearexpats.org/> en <https://twitter.com/wearexpats> (17-04-2018).

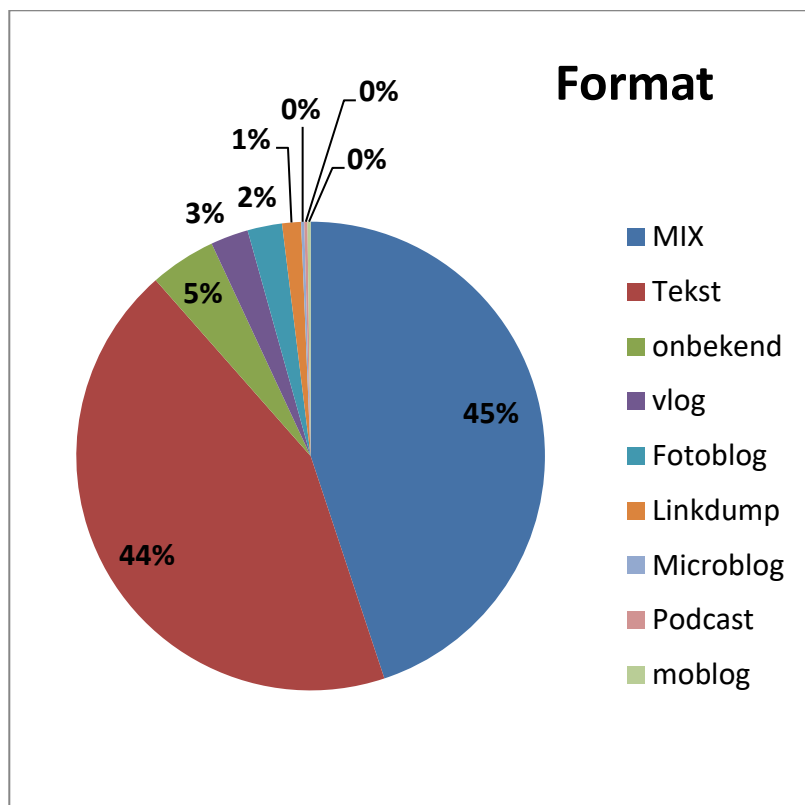
<https://web.archive.org/web/20180528105941/http://wearexpats.org/>

<https://web.archive.org/web/20180528110114/https://twitter.com/wearexpats> (28-05-2018). De twitterlink is door de Wayback Machine om één of andere reden in de Duitse versie gearchiveerd.

websites is deze term overbodig aangezien veel blogs via een mobiel worden bijgehouden.⁷⁴ Omdat de Loglands website het eerste moblog van Nederland is, is gekozen om deze als format toch de tag van 'moblog' te geven.

Op 20-06-2018 waren dit de volgende concrete aantallen⁷⁵:

MIX	207
Tekst	201
Onbekend	21
Vlog	12
Fotoblog	11
Linkdump	6
Microblog, Podcast en Moblog (elk 1)	1



⁷⁴ 'Een moblog is een website of album waarop een gebruiker video's en/of foto's plaatst, dikwijls gepaard gaande met begeleidende tekst en titel, welke gemaakt is met een mobiele telefoon. Deze foto's worden direct op de website geplaatst door deze per mobiele telefoon te verzenden via MMS of e-mail. Bij beide wordt gebruikgemaakt van het GPRS- of UMTS-netwerk van een telecomprovider.' <https://nl.wikipedia.org/wiki/Weblog> (09-05-2018) en <https://nl.wikipedia.org/wiki/Moblog> (12-04-2018). <https://web.archive.org/web/20180528110351/https://nl.wikipedia.org/wiki/Moblog> (28-05-2018).

⁷⁵ Analyses in deze collectiebeschrijving zijn gemaakt op basis van de selectielijst van 20-06-2018. Deze selectielijst wijkt af van de huidige NL-blogsfeer collectielijst.

8.3 Genre

Welke categorieën houden anderen aan?

In bijlage 3 staat in een tabel een overzicht van verschillende genres die de verschillende bronnen onderscheiden. Per boek (en Wikipedia) staat vervolgens aangegeven welke thema's gehanteerd worden of dat er een soortgelijke term wordt gehanteerd.

Indeling naar genres

Deze indeling werkt als volgt: het dikgedrukte thema is het **hoofdthema**. Daaronder staan voorbeelden van mogelijk blogs die hieronder vallen. De **rode** categorieën komen uit de indeling van Martijn Kleppe, de zwarte thema's komen uit één (of meerdere) van de zes bronnen genoemd in bijlage 3.⁷⁶

- **Algemeen**
 - o Omniblogs: blogs met een heleboel onderwerpen

Zoals <http://esthervuijsters.nl>. Dit blog bevat haar persoonlijke lifelog, een giftblog (over cadeautjes), een aantal random blogs ('bloggerie') en dan nog een aantal blogs die door gastschrijvers zijn geschreven (die ook een aantal verschillende onderwerpen hebben).

- o Nicheblogs: blog met één specifiek onderwerp, die niet in een andere categorie valt.
- **Lifelog** (persoonlijk dagboek)

Zoals <http://www.10e.nl> die blogt en plogt (foto blogt) over allerlei elementen van haar persoonlijke leven.

- **Lifestyle**
 - o Design
 - o DIY blog
 - o Foodblog Zoals <http://culunair.nl/>
 - o Reisblogs Zoals <http://letmestayforaday.com/>
 - o Tuinlogs
 - o Cultuur Zoals <https://www.christmaholic.nl/>
 - o Kunst (besprekingen over musea en kunstvoorwerpen)
 Zoals <https://www.digitalekunstkrant.nl/>

⁷⁶ Voor meer informatie over het onderzoek van Martijn Kleppe of welke thema's precies uit welke van de zes genoemde bronnen komen, zie bijlage 3.

- Babysite
- Mamablog Zoals <https://www.ladylemonade.nl/>
- Senioren blogs
- Dieren blogs (katten, honden, etc.)
- **Daten**
- **Gezondheid** Zoals <https://www.fitsociety.nl/>
- **Verslavingen** (gokken, etc.)

- **Mode**
 - Fashion blog (kleding, accessoires, schoenen, etc.)
 - Beauty blogs (make-up, haarverzorgstips, etc.)

- **Politiek weblog**
 - Politieke partijen / politici
 - **Goede doelen**
 - **Religie** Zoals <https://goedgelovig.wordpress.com/> en <http://holtslag.org/>
 - **Overheid**

- **Milieu**
 - Activisten
 - Ecologie Zoals <http://www.interessantetijden.nl/>
 - Klimaat verandering Zoals <https://klimaatverandering.wordpress.com/>
 - **Energie / milieu** (uitleg / bespaartips)

- **Actualiteit** (los van het onderwerp)
 - **Weer & verkeer**
 - Shocklog Zoals <https://www.geenstijl.nl/>
 - Stadsblog (actualiteit van de stad) Zoals <http://www.amsterdamcentraal.nl/>
 - Nieuwslog / liveblogs Zoals <https://www.nu.nl/>

Maar ook blogs die live bijgehouden worden van gebeurtenissen die op het moment gebeuren zoals <http://loglands.nl/>

- **Media recensies**
 - Boekenblogs zoals <http://www.boekenstrijd.nl/>
 - Filmblog
 - Gameblogs
 - Muzieklog Zoals <http://www.metgitarenenzo.nl/>
 - Tv-blog

- Tech blog (over de nieuwste technische snuffjes)
- Mix van de bovenstaande elementen Zoals <http://nl.pcmag.com/>

- **Wetenschappelijk / expertise**

- Blawgs (rechtspraak)
- Programmeerblogs
- Wetenschapsblog

Zoals dit blog door een wetenschapper in de Nieuwe Media & Digitale Cultuur:

<http://www.annehelmond.nl> maar ook bijvoorbeeld <https://www.frankwatching.com/over/> een specialistisch blog over marketing, communicatie, social media en tech. Of <http://astrologieblog.nl/> een blog over astrologie door een astrologe.

- **Educatie**

Deze categorie richt zich op onderwijs, voor zover niet ingedeeld bij de categorie wetenschappelijk/expertise. Het gaat hier met name om blogs in en rond het onderwijs.

- Klasblog
- Blogs als leermiddel
- **Blogs door leerlingen / studenten** Zoals <https://mastersofmedia.hum.uva.nl/>
- Docentenblog

Als docentenblog had ik graag het blog van meester Bart op willen nemen in deze collectie. Hij is echter uitsluitend actief op Facebook en Twitter. Meester Bart is een vrij bekend figuur op het moment. Zijn blog betreft alles over onderwijs vanuit het perspectief van de docent (Engels) die hij is.⁷⁷

- **Hobby**

- Blog over blogs
- Kunst blog (wanneer er kunst wordt geproduceerd door de blogger) zoals <http://www.rikmulder.nl/> zijn fotografie blog. Of <http://www.loods32.com/>
- Afzonderlijke hobby's (bijvoorbeeld sport, voertuigen, creatieve hobby's)

• **Zakelijk blog**

- **Financiële blogs** (bijvoorbeeld blogs over hoe je geld kunt besparen)
- Corporate blog of Businessblog: voor interne / externe bedrijfsvoering
- Commercieel blog (verkoop van producten) Zoals <http://judithvalentijn.nl/>

⁷⁷ <http://www.openschoolgemeenschapbijlmer.nl/meester-bart-is-hit-op-internet/> (04-06-2018). De oudste versie van het artikel staat op 23 augustus 2017. Er staat geen datum bij het oorspronkelijke artikel wanneer het geschreven is. <https://web.archive.org/web/20180604122137/http://www.openschoolgemeenschapbijlmer.nl/meester-bart-is-hit-op-internet/>

- Over je carrière

<http://www.beroepseer.nl/nl/groepsblog> was niet het blog wat ik had verwacht in deze categorie maar toch een voorbeeld van een echt beroeps relateert blog. Dit blog gaat over de werkomstandigheden van arbeidsgroepen. Vakliteratuur valt onder wetenschap/expert.

*Niet ingedeeld*⁷⁸

- Superblogger (term uit Wikipedia)

‘Een verzameling weblogs onder 1 hoofdsite waarin individuele gebruikers een compleet blogstelsel per weblog toegewezen krijgen in plaats van een aangepaste multisite met veelal beperkte mogelijkheden. Een superblogger typeert zich door zeer veel instelmogelijkheden voor complexe blogs met verregerende interactie tussen bloggers onderling en tussen blogger en bezoeker.’⁷⁹

<http://blogonline.blog.nl/over-blognl> is hier denk ik een goed voorbeeld van. Deze overkoepelende blog platformen zijn niet opgenomen in de collectie.

Dit zegt internetdummies (2015) over de superblogger: ‘Er zijn zoveel blogs met zoveel inhoud dat er websites ontstaan waarop het beste materiaal op een plek wordt verzameld. Enkele van de samengestelde websites hebben ladingen auteurs die voor hen schrijven’⁸⁰.

- Themablog: te generieke term van Wikipedia.
- Vrijheid meningsuiting: ook te generiek, van Holthuijsen, in bijna alle blogs zit een mening over van alles en nog wat. En wanneer het over extreme meningen gaat, kan het beter onder politiek of onder shocklog.

Problemen

Voordat ik begon met het indelen van de blogs was ik optimistisch over deze indeling. De werkelijkheid bleek echter ingewikkelder dan gedacht. Blogs veranderen namelijk door de tijd heen regelmatig van onderwerp. Ook bestaan blogs vaak uit posts over vele verschillende onderwerpen die de auteurs toevallig interesseren. De blogs kunnen bovendien allemaal ingedeeld worden onder ‘actualiteit’ omdat vele van hen vaak de actualiteit bespreken. De indeling op genre is dan ook mijn moedige poging om ze toch nog een hoofdthema te geven waar het blog meestal (maar dus niet altijd) over gaat.

Tijdens het vormen van de collectie is de thema indeling licht veranderd: Lifelog stond onder Lifestyle maar heeft zijn eigen kopje gekregen. Hetzelfde geldt voor milieu, wat eerst onder politiek

⁷⁸ Op basis van bijlage 3.

⁷⁹ <https://nl.wikipedia.org/wiki/Weblog> (09-04-2018).

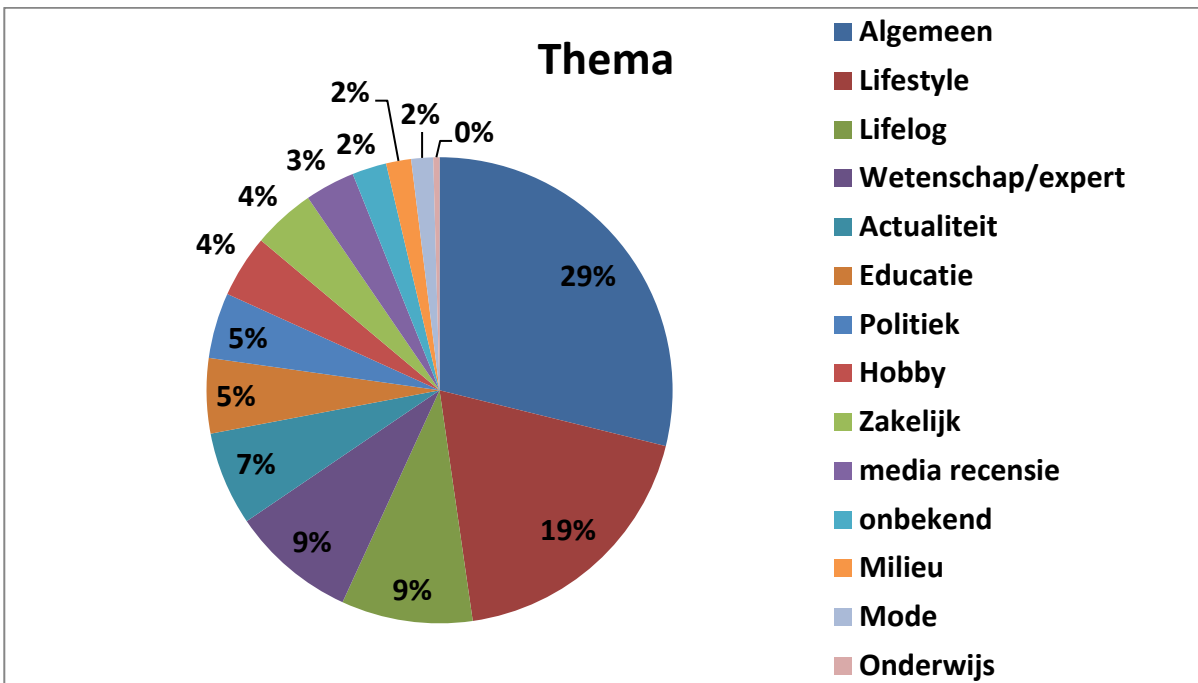
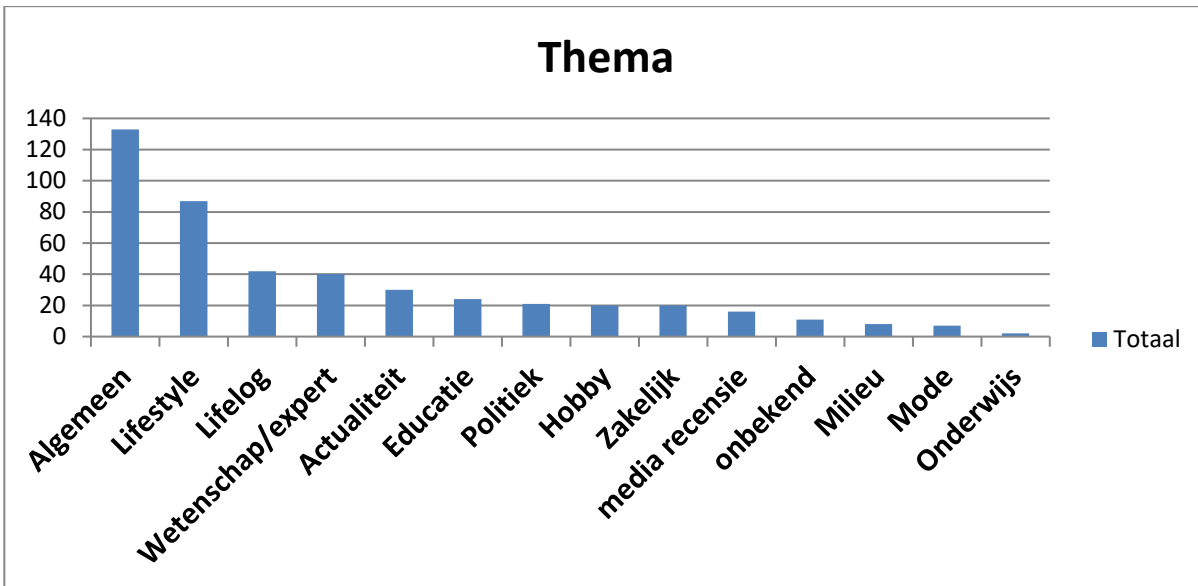
⁸⁰ Levine, J.R. en M. Levine Young, *Internet voor dummies. Dé referentie voor iedereen!* (Amersfoort 2015) 306. Dit staat ook in ongeveer dezelfde bewoording in: Levine, J.R. en M. Levine Young, *Internet voor dummies. Dé referentie voor iedereen!* (Amersfoort 2018).

stond. Liveblogs zijn uiteindelijk onder de actualiteit geplaatst, los van het onderwerp waarover het liveblog gaat. De actualiteit (het live kunnen volgen van iets) is (volgens mij) belangrijker dan het onderwerp/thema.

Statistiek op 20-06-2018⁸¹

Thema	Aantal
Algemeen	133
Lifestyle	87
Lifelog	42
Wetenschap/expert	40
Actualiteit	30
Educatie	24
Politiek	21
Hobby	20
Zakelijk	20
Media recensie	16
Onbekend	11
Milieu	8
Mode	7
Onderwijs	2

⁸¹ Analyses in deze collectiebeschrijving zijn gemaakt op basis van de selectielijst van 20-06-2018. Deze selectielijst wijkt af van de huidige NL-blogosfeer collectielijst.



8.4 Parels & beschrijving

Uitlichting van opvallende delen van de collectie en beschrijving hiervan.

- **Loglands**

Deze website is offline maar was wel geselecteerd voor de collectie NL-blogosfeer. Gelukkig zit hij wel (gefragmenteerd) in de Wayback Machine. Loglands was de eerste moblog (blog bijgehouden met een mobieltje) in Nederland.

- **De 24 grondleggers uit *Bloghelden***

Van de vierentwintig grondleggers uit *Bloghelden* zitten er op dit moment drie van in de collectie. Twee hiervan (Alt en Tonie) zijn onvolledig gearchiveerd. Hier missen belangrijke archiefelementen. L-rs.org is wel gearchiveerd. Hopelijk kunnen in de toekomst de andere websites (via internetarcheologie) toegevoegd worden aan deze collecties wanneer blijkt dat zij aan onze blogvoorwaarden voldoen.

1. E-Wave	13. MisterEdBlog
2. Daily Planet	14. VanTienen.com
3. Het Dagelijks Nieuws	15. Weblogger.nl
4. Prolific (technisch niet mogelijk i.v.m. sociale media verwijzingen) ⁸²	16. ISDIB
5. Smoel.nl	17. Hagenaaars.net
6. Pjoe.net	18. Linked
7. sikkema.net	19. TheWolf
8. MoreWhereThat	20. Stronk
9. Alt0169	21. Cyberramon
10. Joglog	22. L-rs.org
11. Tonie.net	23. Zo.nu/&.dan
12. :=Esc	24. Planeet Eendvogel

- **Letmestayforaday**

Bijzonder blog over de reis van Ramon Stoppelenburg. Hij heeft van mei 2001 tot juli 2003 de wereld rondgereisd zonder geld. Hij werd door mensen die zijn blog volgden uitgenodigd om bij hen te blijven logeren.

Na aanleiding van het archiveren van de website verscheen het volgende artikel in De Stentor (Zwolle): <https://www.destentor.nl/zwolle/koninklijke-bibliotheek-bestempelt-website-oud-zwollenaar-tot-digitaal-erfgoed~accc920d/>

Uit het artikel: „Heel bijzonder”, vindt Stoppelenburg. Hij reageert via skype vanuit Phnomh Penh in Cambodja, waar hij sinds enkele jaren woont. „Wie had dat gedacht toen ik in 2001 met enkele vrienden aan de bar van studentencafé Het Vliegende Paard in Zwolle zes woorden op een viltje krabbelde? Ik hoor het een vriendin nóg zeggen: 'Wie gaat jou nou uitnodigen?'.“⁸³

⁸² Voor meer informatie over (het archiveren van) de Prolific website zie het artikel: Geldermans, I., 2019. We blog: de casus van weblogs in webarchieven. TMG Journal for Media History, 22(1), pp.85–102. DOI: <http://doi.org/10.18146/tmg.438>

⁸³ Geschreven op 02-05-2018 door Janske Mollen.

- **Wij blijven hier!**

Een blog door Nederlandse moslims die niet-Moslims een kijkje willen geven in hun dagelijkse bezigheden en kijk op het leven. Het wordt door meerdere mensen geschreven.

- **Blogwedstrijden**

Dutch Bloggies winnaars: 2001 – 2009

Dutch Bloggie: 'Blogger van het decennium': 2010

- Winnaars & genomineerden

Blog Parels: 2010, 2011 en 2013

Blogger des Vaderlands: 2015

- Winnaars, jury & genomineerden

Deze blogwedstrijden geven min of meer van 2001 tot 2015 een beeld van de blogs die destijds gewaardeerd werden als prijswaardig. Een opmerking hierbij is wel dat de Dutch Bloggies vaak ook tot een storm van kritiek leidde en dat GeenStijl er bijvoorbeeld in 2004 in slaagde alle publieksprijzen te kapen door hun leger van volgers.⁸⁴ Dit zegt echter ook iets over de kracht van shocklog GeenStijl.

⁸⁴ *Bloghelden*, 153 – 154.

9 Linkanalyse

9.1 Hyphe linkanalyse

Hyphe analyse met de hyperlinks dataset van 20-06-2018.⁸⁵

Meest gelinkt:

Top 10 'discovered' websites:

	Website	Indegree
1.	Youtube.com	192
2.	Wikipedia.org	152
3.	Instagram	119
4.	Google.com	114
5.	Google.nl	71
6.	Nos.nl	70
7.	Wordpress.org	68
8.	Twitter.com	67
9.	Bol.com	59
10.	Flickr.com	58

Top 10 'In' websites:

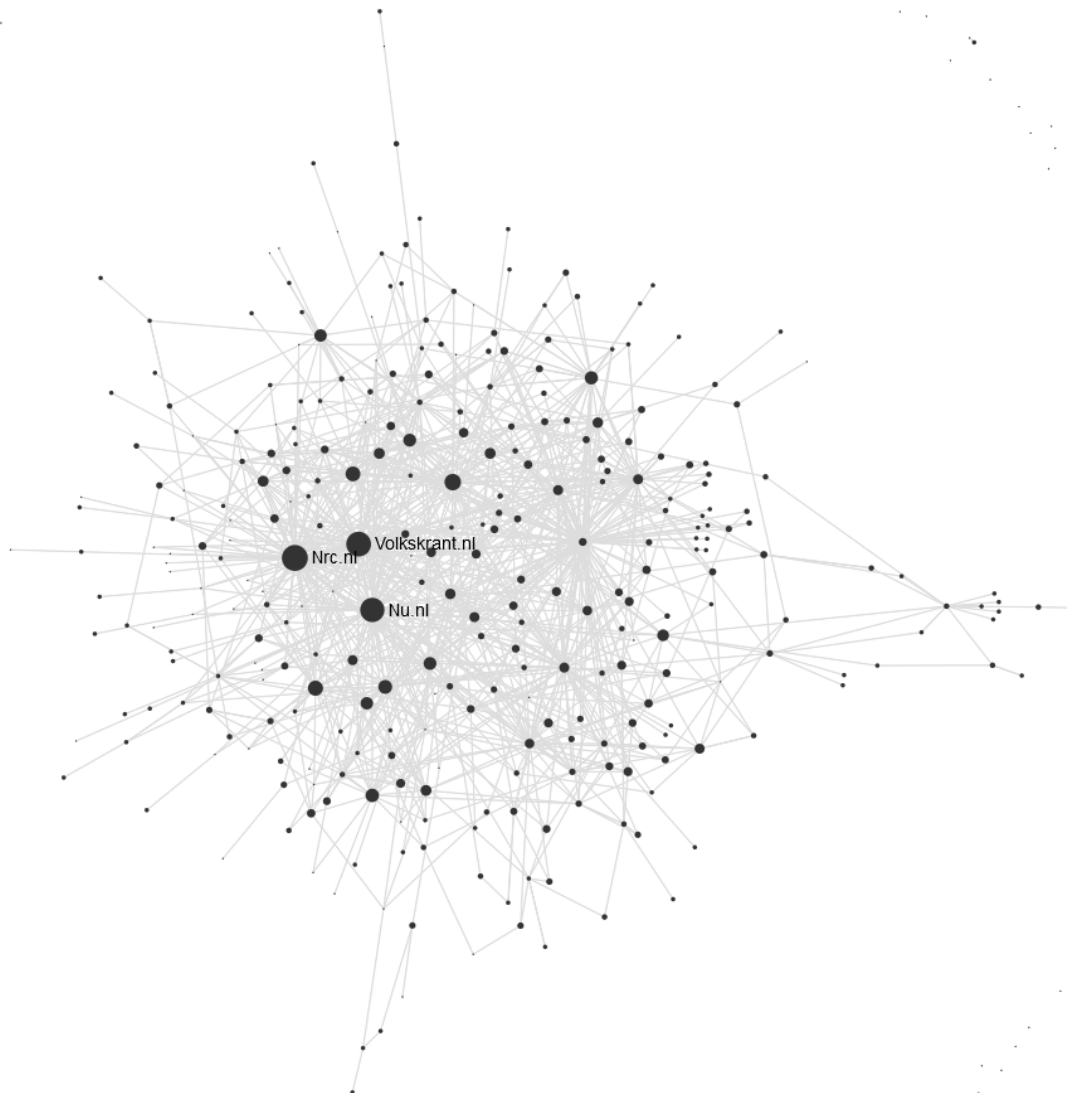
	Website	Indegree
1.	Nrc.nl	75
2.	Volkskrant.nl	68
3.	Nu.nl	65
4.	Sargasso.nl	28
5.	Bnr.nl	23
6.	Geenstijl.nl	22
7.	Marketingfacts.nl	19

⁸⁵ Analyses in deze collectiebeschrijving zijn gemaakt op basis van de selectielijst van 20-06-2018. Deze selectielijst wijkt af van de huidige NL-blogosfeer collectielijst.

	Website	Indegree
8.	Tweakers.net	18
9.	ximaar.Wordpress.com	17
10.	Denieuwereporter.nl	16
	Metronieuws.nl	16

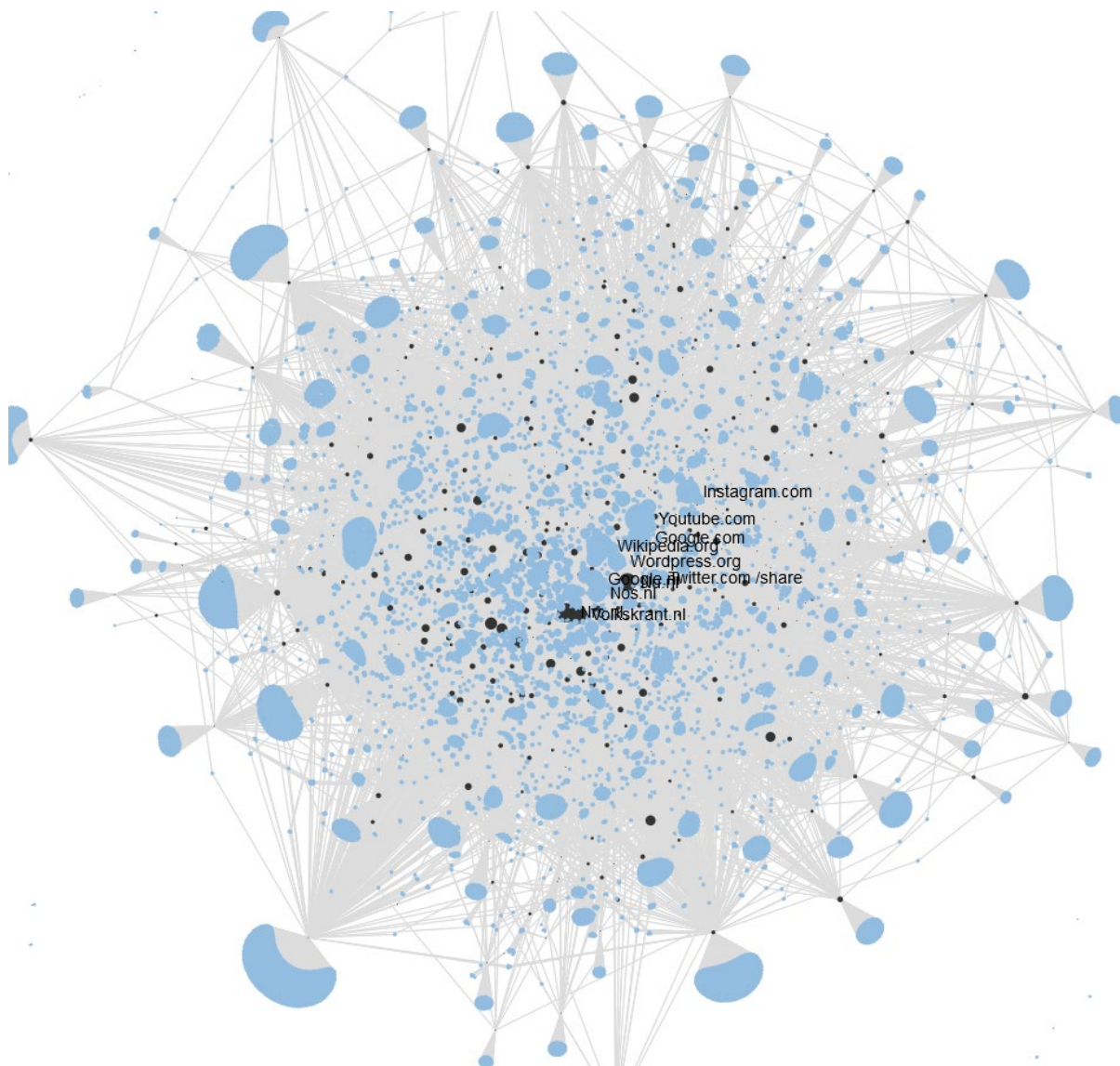
Visualisatie alle 'IN' records

(records die zijn ingevoerd / de links uit de database)



'IN' records en 'Discoverd' records samen gevisualiseerd:

Discovered records = websites gevonden door de linkanalyse. Websites zaten dus niet oorspronkelijk in database.

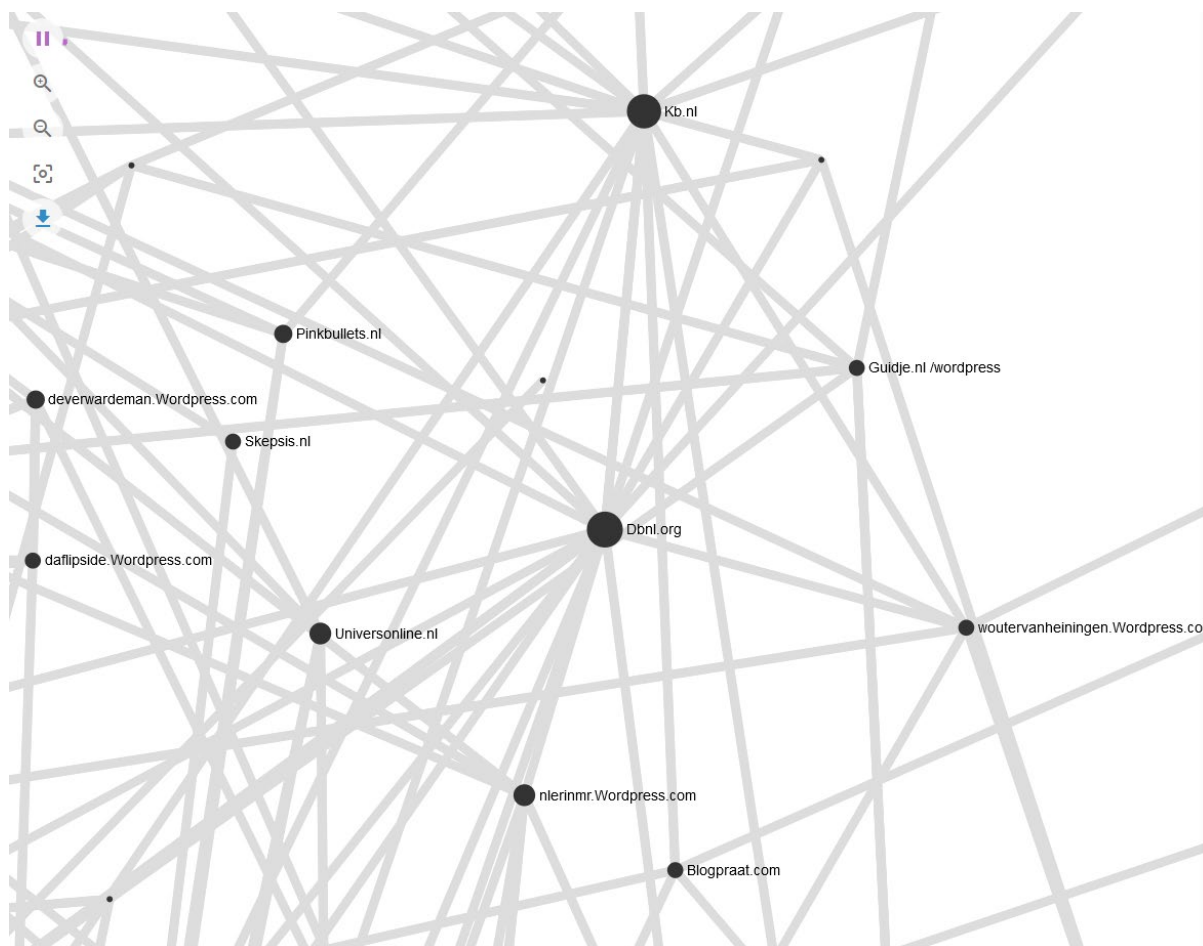


Linkanalyse positie KB websites:

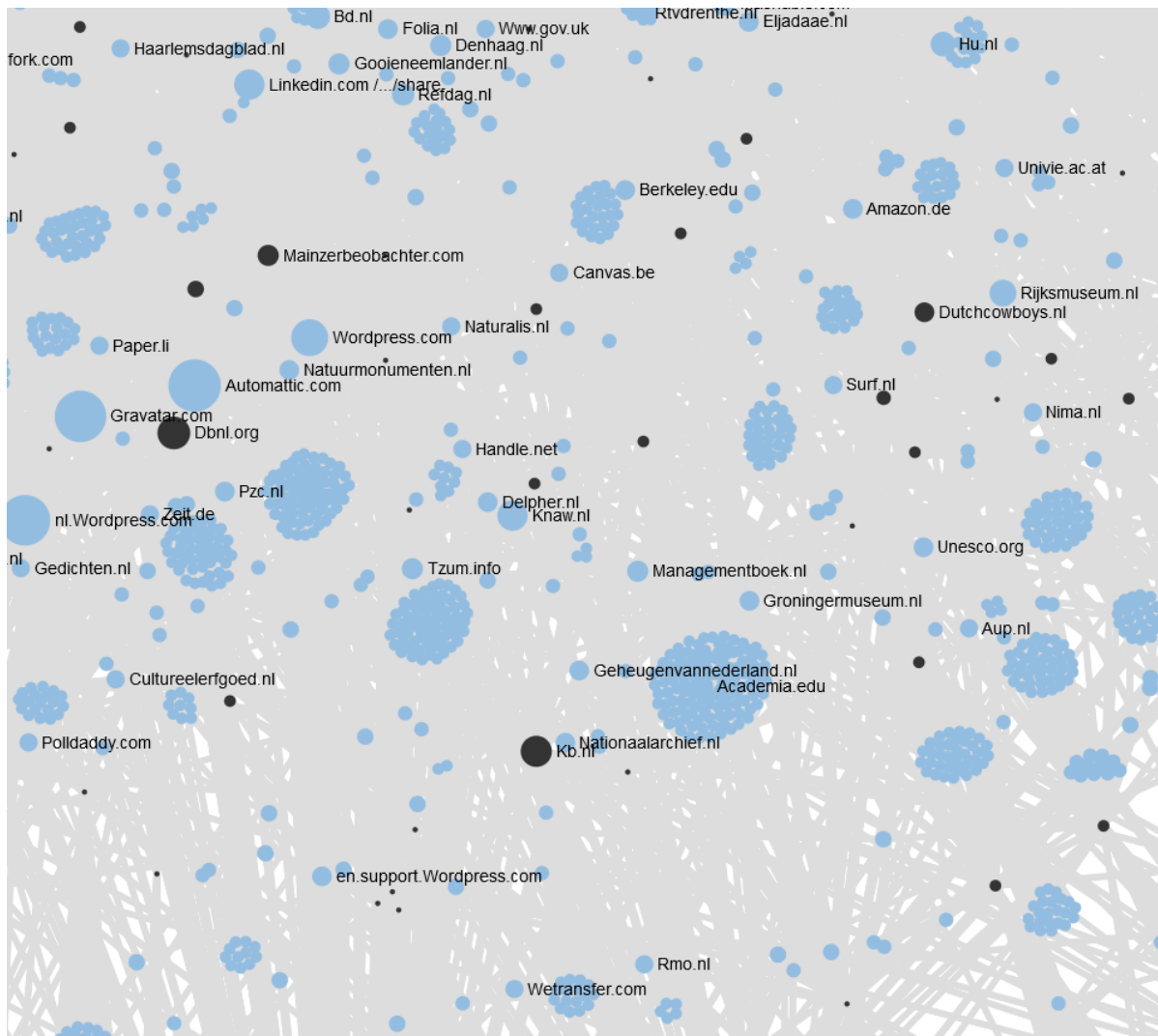
DBNL	indegree = 18	Plaats 71
KB	Indegree = 16	Plaats 90
Delpher	indegree = 5	Plaats 385

Gesorteerd op indegree en daarna op alfabet.

IN records visualisatie:



In records & Discovered records visualisatie:



Websites van de KB liggen bij elkaar omgeven door andere wetenschappelijke instellingen zoals Unesco, Groningen Museum, Rijksmuseum, Naturalis, Cultureel Erfgoed, Nationaal Archief, Geheugen van Nederland en Natuurmonumenten.

10 Datasets

Wij hebben er voor gekozen om voor onze collectie blogs kwalitatief te beoordelen en aan de hand daarvan toe te voegen aan de collectie. Omdat het Nederlandse webdomein zo groot is, zijn we niet toegekomen aan alle blogs. Voor toekomstig onderzoek zijn daarom drie bonuslijsten beschikbaar met websites waarvan wij sterk vermoeden dat het blogs zijn maar die nog niet kwalitatief beoordeeld zijn. Het kan zijn dat deze lijsten deels overlappen met elkaar en met de collectie. Het gaat om de volgende lijsten:

10.1 Helmond lijst (bonuslijst Helmond Dutch blogosphere schoon en ontdebeld)

Dit is de dataset die Anne Helmond en Esther Weltevrede hebben gebruikt voor hun onderzoek in 2012.⁸⁶ Ik heb deze opgeschoond en ontdebeld (van 2500 naar 1500). Er zaten (vermoed ik) veel dubbele records in vanwege een spatie of slash teken (/) aan het eind van sommige internet adressen en het verschil in `http://www.` en alleen `http://` aan het begin van sommige hyperlinks. De lijst zal nog een keer grondiger ontdebeld moeten worden. Ik heb gemarkeerd wat ik kon vinden maar ik vermoed dat er nog wel wat dubbele records in zitten.

In deze lijst zit de loglijst van Jan-Willem Hiddink en Robert Reinder waar Meeuwse over schrijft in *Bloghelden* (bladzijde 57).⁸⁷ De loglijst in de Helmond lijst bestaat uit 566 records en is een interessant snapshot van de blogosfeer in Nederland.

Extra kolom⁸⁸:

- Schone URL URL zonder `www`, `http`, aan het begin of slash of spatie aan het einde.

10.2 Ongeldige links (bonuslijst ongeldige links)

Een zelf samengestelde lijst met websites die niet langer online zijn of websites waarvan ik wel weet dat het een blog is (bijvoorbeeld door literatuur onderzoek) maar er geen internetadres te vinden is.⁸⁹ Deze links zijn niet opgenomen in de collectie.

Geen extra kolommen.

⁸⁶ Weltevrede, E. and Helmond A. (2012) Where do bloggers blog? Platform transitions within the historical Dutch blogosphere. *First Monday* 17(2). DOI: [10.5210/fm.v17i2.3775](https://doi.org/10.5210/fm.v17i2.3775).

⁸⁷ Hier is een versie van een loglijst waar Robert Reinder aan meewerkt, te zien in het Internet Archive: <http://web.archive.org/web/20020724165825/http://www.nederhoed.com/~jeewee> Het is mij onduidelijk of het de lijst is die Anne Helmond en Esther Weltevrede hebben gebruikt.

⁸⁸ Sommige bonuslijsten bevatten extra kolommen met aanvullende informatie. Deze kolommen zijn ingevoegd aan het einde om de lijsten zo identiek mogelijk te houden. Per bonuslijst staat aangegeven welke kolommen zijn toegevoegd bovenop de kolommen genoemd in paragraaf 8.1.

⁸⁹ Deze lijst is gelijktijdig gemaakt met de NL-blogosfeer selectielijst in april/mei 2018.

10.3 KB blogs (bonuslijst KB blogs)

Binnen de al gearchiveerde websites van de Koninklijke Bibliotheek zitten ook al blogs. De weblogs zijn gevonden via zoekopdrachten binnen de kwartaaluitdraai. De versie die gebruikt is om deze lijst te maken is de uitdraai van 31-03-2018.

- Zoekstring: 'blog'
- Zoekstring: 'wordpress'

Bij de Wordpress websites is nog wel de volgende kanttekening te maken: Wordpress websites worden veel gebruikt voor blogs, maar ook voor zakelijke websites van winkels. Voor het gebruik van deze zoekterm is het zeer belangrijk om goed te kijken of de websites wel blogs zijn. Een filtering of een Wordpress website een blog is of iets anders heeft op deze lijst nog niet plaatsgevonden.

- Zoekstring: 'log'

Deze zoekstring is als volgt opgebouwd.

1. Zoek op 'log'
2. Filter de hits met 'oorlog', 'blog' en 'logisch' eruit
3. Filter 'internetarcheologie' hits eruit
4. Handmatig de overgebleven lijst globaal doorkijken op opvallende verkeerde hits

Voor gebruik zal, net als bij de Wordpress lijst, nog goed gekeken moeten worden of deze websites daadwerkelijk blogs zijn.

Extra kolommen

- UNESCO De unesco indeling die de KB geeft aan een website tijdens het invoeren in de WCT.

Kleurcodering

- Groen: vrij zeker dat het blogs zijn (zoekterm blog)
- Wordpress (geel) & log (rood)
 - (onzekerder of het ook wel echt blogs zijn).

11 Status geselecteerde websites per 31-08-2020

31 augustus 2020

461 websites geselecteerd voor de NL-blogosfeer⁹⁰

336 websites in WCT⁹¹

177 websites tijdens invoerproces afgekeurd of toestemming geweigerd door website eigenaar.

20 juni 2018

461 websites geselecteerd voor de NL-blogosfeer

182 websites in de WCT

119 websites tijdens invoerproces afgekeurd of toestemming geweigerd door website eigenaar.

160 websites nog niet behandeld waarvan er 160 al kwalitatief beoordeeld zijn & klaar zijn voor invoer in de WCT.

31 mei 2018

331 websites geselecteerd voor de NL-blogosfeer

104 websites in de WCT

65 websites tijdens invoerproces afgekeurd of toestemming geweigerd door website eigenaar.

162 websites nog niet behandeld waarvan er 85 al kwalitatief beoordeeld zijn & klaar zijn voor invoer in de WCT.

⁹⁰ Dit is de selectielijst van 20-06-2018. Deze selectielijst wijkt af van de huidige NL-blogosfeer collectielijst.

⁹¹ In de periode na 20 juni 2018 hebben ook websites in het webarchief die niet op de 'websites geselecteerd voor de NL-blogosfeer' lijst stonden het collectielabel NL-blogosfeer gekregen wanneer de website aan de selectiecriteria voldeed.

12 Conclusie

12.1 Algemene conclusie

Er zijn ontzettend veel soorten blogs te vinden op allerlei websites in allerlei soorten en maten. Wat op dit moment (4 juni 2018) nog wel mist in de collectie zijn de podcasts, de microblogs en de (pure) modeblogs. Zeker van de modeblogs had ik er veel meer verwacht. Ik heb hier niet verder aan kunnen werken vanwege tijdsgebrek.

Weblogs en nieuws zijn sterk met elkaar verbonden. Nieuwskranten (en andere media) hebben veel stijlfiguren overgenomen van de oerbloggers (zoals in de Enschede case in Bijlage 1 (17.1) te zien is). Zou het zo kunnen zijn dat door de snel reagerende bloggers de mainstream media moest volgen? Bloggers reageerden snel op de actualiteit. Lijken weblogs en nieuws daarom zo ontzettend op elkaar?

Mensen noemen hun blog niet altijd een blog ondanks dat een website wel alle kenmerken heeft van een blog. Wanneer mensen hun blog een blog noemen heb ik ze meestal opgenomen. Maar ik heb ook vaak websites opgenomen die zichzelf als 'magazine' omschrijven. Of andere termen. Het zou intrigerend kunnen zijn om te kijken of dit vooral op mannelijk publiek gerichte websites zijn.

Een voorbeeld van zo een website is de <https://www.digitalekunstkrant.nl/>

Korte analyse van de digitale kunstkrant toont dat het bijna alle elementen van een blog bevat:

- ✓ Een reeks afzonderlijke berichten (posts)
- ✓ Waarbij een omgekeerde chronologische volgorde wordt aangehouden
- ✓ Er verschijnen regelmatig nieuwe berichten: het liefste dagelijks of wekelijks.
- ✓ Een blog staat online op een aparte, eigen, website of is een onderdeel van een website.
- ✓ Een blog kan ook een combinatie zijn van tekst, afbeeldingen, geluid, (interne en externe) links en video's.
- ✓ Blogs kunnen openbaar of gesloten zijn.
- ✓ Iedereen kan bloggen: van amateurs tot professionals. Je kunt er zelfs je beroep van maken.
- ✓ Een blog kan door één of meerdere personen geschreven worden.
- ✓ Er is iemand (of meerdere personen) verantwoordelijk voor de inhoud. Niet iedereen kan zomaar berichten (posts) plaatsen op een blog.

Veel voorkomende elementen (maar geen verplichte onderdelen)

- ✓ Bezoekers kunnen meestal reageren op een blog
- ✓ Een blog kan een archief hebben waarin alle bijdragen bewaard worden
- ✓ Tijd en datum staan meestal bij de post genoemd

Het blog als primair bronmateriaal

Blogs zijn een geweldige bron voor toekomstig onderzoek: het zijn de dagboeken van deze tijd.

- *Bottom up* onderzoek: onderzoek naar het volk: hun reacties op actualiteit en zaken die hen in het dagelijks leven bezighielden. (Zoals bijvoorbeeld de vuurwerkramp in Enschede).⁹²
- De reactiemogelijkheid in blogs toont eventuele discussie over een bericht of thema. Het laat wellicht verschillende kanten en meningen zien betreft een thema / gebeurtenis.

Het gapende gat: sociale media

Social media & Youtube missen ontzettend in deze collectie. Zij missen als onderdeel van een blog (verwijzingen naar filmpjes bijvoorbeeld) en zij missen als complete microblog (Meester Bart en Prolific bijvoorbeeld). Doordat we deze data niet kunnen archiveren missen (in sommige gevallen) cruciale onderdelen van een blog of complete blogs.

Op 30 mei is er een kleine linkanalyse gedaan door middel van het programma Hyphe (slechts één laag diep) met 327 blogs. Hieruit zijn 13.543 website/verwijzingen naar externe websites gevonden. Social media scoren relatief hoog in deze analyse betreft indegree (meest gelinkte websites). Hieronder staat de top 10 meest gelinkte websites⁹³:

	Link	Indegree
1	Twitter.com	196
2	Facebook.com	181
3	Youtube.com	133
4	Google.com	113
5	Wordpress.com	109
6	Wikipedia.org	104
7	Instagram.com	85
8	Linkedin.com	83
9	Pinterest.com Volkskrant.nl Nos.nl	55
10	Google.nl	50

⁹² Geldermans, Iris, 'NL-blogosfeer: Burgers verslaan een ramp' website *Koninklijke Bibliotheek | Nationale Bibliotheek* (27-08-2018) <https://www.kb.nl/blogs/duurzame-toegang/nl-blogosfeer-burgers-verslaan-een-ramp>

⁹³ Analyses in deze collectiebeschrijving zijn gemaakt op basis van de selectielijst van 20-06-2018. Deze selectielijst wijkt af van de huidige NL-blogosfeer collectielijst.

Hieronder de meest genoemde social media hits.

Link	Indegree
Twitter.com	196
Facebook.com	181
Youtube.com	133
Instagram.com	85
Linkedin.com	83
Pinterest.com	55
Flickr.com	44
Youtu.be	34
Tumblr.com	32
Vimeo.com	27
Snapchat.com	5

Hoewel bloggen op het moment nog redelijk stabiel aanwezig is, kan het zo gebeuren dat social media die rol compleet gaat overnemen. Zo is Instagram ontzettend in opkomst. Overigens is de blog webpagina vaak ook weer een verzamelplaats voor een professionele blogger waar alle social media kanalen samenkomen. Bezoekers krijgen hier een goed overzicht op welke platformen de blogger allemaal actief is. Het is daarom jammer dat sociale media niet kunnen worden meegearchiveerd. Het is vaak (zeker bij moderne blogs en zeker bij vlogs) een belangrijke doorverwijzing en onderdeel.

12.2 Aanbevelingen

1. Social media

Hoewel ik weet dat het op het moment niet mogelijk is, is het toch mijn aanbeveling om een manier te vinden om sociale media te archiveren. Niet alleen in de NL-blogosfeer zijn social media onmisbaar, ook een groot deel van ons leven speelt zich op deze platformen af. Door de afwezigheid van sociale media in webarchieven mist de onderzoeker een cruciaal historisch overzicht van onze samenleving.

2. Excel

Er is veel mogelijk met Excel. Wellicht is het een idee om meer kennis op te doen over:

- Onderzoek in Excel (simpele statistieken)
- Preprocessing in Excel (data opschonen in Excel)
- Database & Excel:

- wat kan een onderzoeker al aflezen uit de gegevens van Excel / wat wil een onderzoeker kunnen aflezen in een Excel.
- Standaardisering webcollecties in de KB (in Excel)
 - Welke informatie moet in een webcollectie beschikbaar zijn
 - Welke informatie mogen collectievormers zelf toevoegen
 - 'Merge punt': de naam

Over het 'merge punt' heb ik al eerder gesproken. Mochten onderzoekers in de toekomst van de verschillende webcollecties van de KB één grote collectie willen maken waarbij alle metadata per record beschikbaar is, dan moet er nagedacht worden over (en nu al rekening gehouden worden met) het 'merge punt'. Het punt waarop de records aan elkaar gebreed kunnen worden.

De collectiespecialist van het webarchief had hierover een goede suggestie. Een record in de WCT van de KB heeft een uniek nummer. Dat is eenvoudig aan de Excelbestanden toe te voegen. Wanneer in alle speciale collecties consequent dit unieke ID nummer wordt gebruikt is dit een uitstekend 'merge punt' waarmee alle informatie aan elkaar gebonden kan worden in één database.

13 English summary

Written on the 22th of October 2020

In 2018 I, Iris Geldermans, worked as an intern for the Koninklijke Bibliotheek to form a collection of Dutch weblogs for the webarchive named the NL-blogosfeer. During this internship I discovered that weblogs are important primary sources for historical research because they place an event in the context of its time. Bloggers continuously write about what is happening right now. Lots (but not all!) of the weblogs are a bottom-up source: they give us a glimpse of everyday life and opinions of normal people. Weblogs show how people reacted on current affairs such as elections, festivals, #Me2, changing fashion trends, 9/11, MH17 and the firework disaster in Enschede. Blogs, however, are also vulnerable. Like so many websites, weblogs disappear from the internet, and with their disappearance we lose their archives often covering many years, the reactions placed under blogposts and their discussions on (now historical) events.

To prevent loss of the rich Dutch blogging culture, the Koninklijke Bibliotheek decided to form a collection of blogs under the name NL-blogosfeer. This collection has as goal to give a representative image of the Dutch blogging culture and is meant as a primary source for reflection on Dutch digital culture. Websites are also selected because of their national cultural historic importance. Another reason to start this collection was het research *where do bloggers blog? Platform transitions within the historical Dutch blogosphere* by Esther Weltevrede and Anne Helmond.⁹⁴ In this paper Weltevrede and Helmond analyze the Dutch blogosfeer on the basis of expert lists found in literature (like *Bloghelden* by Frank Meeuwssen⁹⁵) and through interviews with bloggers that were involved with the Dutch Bloggie awards.

⁹⁴ Weltevrede, E. and Helmond A. (2012) Where do bloggers blog? Platform transitions within the historical Dutch blogosphere. *First Monday* 17(2). DOI: [10.5210/fm.v17i2.3775](https://doi.org/10.5210/fm.v17i2.3775).

⁹⁵ You can view and download the PDF of the book (written in Dutch) on this website: <http://bloghelden.nl/>

14 Verder onderzoek

14.1 Waar zouden nog meer blogs gevonden kunnen worden?

- Er zou gekeken kunnen worden naar de lijst van blogs wereldwijd en dan het .nl domein eruit filteren.
- www.startpagina.nl bekijken in het Internet Archive. Daar kun je wellicht bekijken in verschillende jaren welke blogs zij destijds benoemde op de startpagina. In die tijd was startpagina.nl een populaire website.
- Op <https://weblog.startpagina.nl/> kunnen blogs van dit moment gevonden worden
- Hyped.nl top 30 weblogs die het meeste waard zijn:
 - o 2009, februari
https://web.archive.org/web/20090312023352/http://www.hyped.nl:80/details/20090204_weblog_top_30_geenstijlnl_14_miljoen_waard
 - o 2009, september
<https://web.archive.org/web/20100308115049/http://www.hyped.nl:80/pages/weblogwaarde>

14.2 Onderzoeksideeën

- Linkanalyse: welke websites worden gevonden en waar verwijzen zij naar door? Hoe verbonden is de blogosfeer?
- *Close reading*: hoe reageerden bloggers op bepaalde gebeurtenissen zoals:
 - o De verkiezing van Trump, de aardbevingsschade in Groningen, verkiezingen, Referenda
 - o Nederlandse verkiezingsuitslagen, festivals, de dood van Pim Fortuyn, #Me2, veranderende modetrends, de reactie op 9/11 en de MH17.

2020

Vanaf 2020 begint het webarchief interessant te worden voor '20 jarige herdenkingen' (*close reading* 'hoe reageert het publiek op' & *distant reading analyses*).

2000:

- Vuurwerkramp Enschede
- Brand in Volendams café

2001:

- Engelstalige Wikipedia gaat van start (linkanalyse gebruik Wikipedia door NL).
- 9/11 & begin van de strijd tegen het terrorisme

- Oprichting Leefbaar Nederland: Pim Fortuyn.

2002:

- Invoer Euro / verlies Gulden
- Huwelijk Willem-Alexander & Maxima
- Pim Fortuyn: Lijst Pim Fortuyn, zijn moord en de daarop volgende verkiezingen
- NIOD & Srebrenica rapport / val kabinet Kok II
- Val Balkenende I

2003:

- De eerste kapsalon (het gerecht) in Nederland
- Vredesdemonstratie tegen oorlog in Irak (wereldwijd, 15 februari)
 - Grootste demonstratie sinds de Vietnamoorlog / wat is de rol van internet?
 - In communicatie / oproep tot het protest
 - In verbondenheid
 - Wordt er veel over gepost?
- Internationale wiewerunie voert helmplicht in na aanleiding van de dood van Andrei Kivilev
- Opening Westerscheldetunnel
- Internationaal strafhof treed in werking voor oorlogsmisdaden
- Air France neemt KLM over

2004:

- Theo van Gogh wordt vermoord

15 Literatuur

15.1 Literatuur

- Bekker, C., *De blog bijbel. Voor beauty chicks & shopaholics* (Maartensdijk 2014).
- Boer, T. de, *Van Dale Pocketwoordenboek Nederlands* (Utrecht 2013).
- Boon, T den, Hendrickx, R. en N. van der Sijs, *Van Dale. Groot woordenboek van de Nederlandse taal* (Utrecht/Antwerpen 2005).
- Duval, S., *How blogs work. Tips & tricks voor beginnende en ervaren bloggers* (Tielt 2014).
- Deriemaeker, K., *Blog boek. Tips & inspiratie voor meer blog- en vlogplezier* (Antwerpen 2017).
- Hofstede, M., *Het zoekboek voor het web* (Rijswijk 2001).
- Holthuijsen, D., *Weblogs. Uw persoonlijke blog op het web!* (Haarlem 2005).
- Kleppe, M. en Otte M., Analysing and understanding news consumption patterns by tracking online user behaviour with a multimodal research design, *Digital Scholarship in the Humanities*, Volume 32, Issue suppl_2, December 2017, Pages ii158–ii170. DOI: <https://doi.org/10.1093/llc/fqx030>.
- Meeuwssen, F., *Bloghelden* (2010).
- Meindersma, C., *Wetboek voor bloggers. De (on)geschreven regels voor bloggers, journalisten en content creators* (Den Haag 2016).
- Weltevrede, E. and Helmond A. (2012) Where do bloggers blog? Platform transitions within the historical Dutch blogosphere. *First Monday* 17(2). DOI: <https://doi.org/10.5210/fm.v17i2.3775>.

Internet dummies

- Levine, J.R. en C. Baroudi en M. Levine Young, *Internet voor dummies. Dé referentie voor iedereen!* (Amsterdam 1997).
- Levine, J.R. en C. Baroudi en M. Levine Young, *Internet voor dummies. Dé referentie voor iedereen!* (Amsterdam 2000).
- Levine, J.R. en C. Baroudi en M. Levine Young, *Internet voor dummies. Dé referentie voor iedereen!* (Amsterdam 2002).
- Levine, J.R. en C. Baroudi en M. Levine Young, *Internet en e-mail voor dummies. Dé referentie voor iedereen!* (Amsterdam 2005).
- Levine, J.R. en C. Baroudi en M. Levine Young, *Internet en e-mail voor dummies. Dé referentie voor iedereen!* (Amsterdam 2007).
- Levine, J.R. en M. Levine Young, *Internet voor dummies. Dé referentie voor iedereen!* (Amersfoort 2015).
- Levine, J.R. en M. Levine Young, *Internet voor dummies. Dé referentie voor iedereen!* (Amersfoort 2018).

15.2 Websites

Naam website	link	WayBack Link Internet Archive	Datum geraadpleegd	Datum Wayback versie
Van Dale website	http://www.vandale.nl/gratis-woordenboek/nederlands/betenis/weblog	https://web.archive.org/web/20180524123949/https://www.vandale.nl/gratis-woordenboek/nederlands/betenis/weblog	16-04-2018	24-05-2018
We Are Xpats	http://wearexpats.org/	https://web.archive.org/web/20180528105941/http://wearexpats.org/	17-04-2018	28-05-2018
We Are Xpats twitter	https://twitter.com/wearexpats	https://web.archive.org/web/20180528110114/https://twitter.com/wearexpats (Duitse versie maar berichten gewoon in het Engels)	17-04-2018	28-05-2018
Shine tool	https://www.webarchive.org.uk/shine	https://web.archive.org/web/20180524123621/https://www.webarchive.org.uk/shine	17-04-2018	24-05-2018
Wikipedia - moblog	https://nl.wikipedia.org/wiki/Moblog	https://web.archive.org/web/20180528110351/https://nl.wikipedia.org/wiki/Moblog	12-04-2018	28-05-2018
Wikipedia - weblog	https://nl.wikipedia.org/wiki/Weblog	https://web.archive.org/web/20180524123500/https://nl.wikipedia.org/wiki/Weblog	09-04-2018	24-05-2018
Venturo weblog	X	https://web.archive.org/web/20130508153845/http://venturo.weblog.nl/	16-05-2018	08-05-2013
Loglands moblog	X	https://web.archive.org/web/20070811060058/http://www.loglands.nl	16-05-2018 08-05-2018	11-08-2007
Bloghelden	http://bloghelden.nl/	https://web.archive.org/web/20180531141015/http://bloghelden.nl/	31-05-2018	31-05-2018
PDF <i>Bloghelden</i> boek		https://web.archive.org/web/20180531141209/https://raw.githubusercontent.com/frankmeeuwse/n/Bloghelden/master/bloghelden-boekversie.pdf	31-05-2018	31-05-2018
Google Trends weblog/blog	https://trends.google.nl/trends/explore?date=all&geo=NL&q=blog.weblog	X	17-05-2018	Staat niet in Internet Archive
Blogonline		https://web.archive.org/web/20180531074216/http://blogonline.blog.nl/over-blognl		31-05-2018
Alleweblogs		https://web.archive.org/web/20180531074431/https://www.alleweblogs.nl/		31-05-2018
Vinex resultaten	https://vinex.nl/resultaten/	https://web.archive.org/web/20180528082940/https://vinex.nl/resultaten/	28-05-2018	28-05-2018
Blogger des Vaderlands	http://hetblogbal.blogspot.nl/p/blogger-des-vaderlands.html	https://web.archive.org/web/20180528083552/http://hetblogbal.blogspot.nl/p/blogger-des-vaderlands.html	09-05-2018	28-05-2018
Peterspagina	http://www.peterspagina.nl/blogparel/ http://www.peterspagina.nl/blogparel_2011/		09-05-2018	

Naam website	link	WayBack Link Internet Archive	Datum geraadpleegd	Datum Wayback versie
Blogparel wedstrijd	http://www.blogparel.nl/		09-05-2018	
Tonie.net	X	https://web.archive.org/web/20010830100146/http://www.tonie.net:80/	30-05-2018	30-08-2001
Planet Internet	X	https://web.archive.org/web/20011115044647/http://web.planet.nl:80/	30-05-2018	15-11-2001
ESC (de digitale stad)	X	https://web.archive.org/web/20010528184124/http://esc.dds.nl:80/	30-05-2018	28-05-2001
Alt0169	X	https://web.archive.org/web/20010421110602/http://www.alt0169.com:80/	30-05-2018	21-04-2001
Twente.nl	X	https://web.archive.org/web/20000302143536/http://www.twente.nl:80/	30-05-2018	02-03-2000
Lines en Splines	http://www.linesandsplines.com/ (Nu: Duits casino!)	PAS OP MET HET OPENEN VAN DE LINK	31-05-2018	Niet van toepassing
Loglijst Robert Reinder	X	http://web.archive.org/web/20020724165825/http://www.nederhored.com/~jeewee	Ergens in mei 2018.	24-07-2002
SIDN : Whois webtool	https://www.sidn.nl/whois	X	28-05-2018	Niet mogelijk
Duurzame Brabander post 1	http://www.duurzamebrabanders.nl/blog/2011/01/maak-een-mooie-backup-op-een-gemakkelijke-manier/	https://web.archive.org/web/20110611093652/http://www.duurzamebrabanders.nl/blog/2011/01/maak-een-mooie-backup-op-een-gemakkelijke-manier/	11-06-2018	11-06-2018
Duurzame Brabander post 2	http://www.duurzamebrabanders.nl/blog/2011/01/zijn-er-nog-duurzame-bloggers-bij-het-vkblog/	https://web.archive.org/web/20110611093442/http://www.duurzamebrabanders.nl/blog/2011/01/zijn-er-nog-duurzame-bloggers-bij-het-vkblog/	11-06-2018	11-06-2018
Dutch Mom Blog Awards	X	https://web.archive.org/web/20110412130350/http://www.momblogawards.nl/	22-12-2020	12-04-2018

15.3 Artikelen op websites (met Internet Archive Wayback link waar mogelijk).

<https://web.archive.org/web/20130416072539/http://incredibleadventure.nl/2013/01/sanoma-stopt-met-weblog-nl-een-terugblik-met-de-oprichter/> (16-05-2018).

<http://webwereld.nl/e-commerce/26807-web-log-nl-begroet-100-000ste-weblog> (16-05-2018).
<https://web.archive.org/web/20180524114958/http://webwereld.nl/e-commerce/26807-web-log-nl-begroet-100-000ste-weblog> (24-05-2018).

<http://webwereld.nl/e-commerce/231-de-8-maanden-migratiemalaise-van-web-log-nl> (16-05-2018).
<https://web.archive.org/web/20180524115304/http://webwereld.nl/e-commerce/231-de-8-maanden-migratiemalaise-van-web-log-nl> (24-05-2018).

<http://webwereld.nl/e-commerce/28258-volkskrant-geeft-lezers-een-stem-via-nieuw-blog> (16-05-2018).

<https://web.archive.org/web/20180524115401/http://webwereld.nl/e-commerce/28258-volkskrant-geeft-lezers-een-stem-via-nieuw-blog> (24-05-2018).

<http://webwereld.nl/development/57010-volkskrant-blogs-al-maanden-offline-door-web-log-nl> (16-05-2018).

<https://web.archive.org/web/20180524115459/http://webwereld.nl/development/57010-volkskrant-blogs-al-maanden-offline-door-web-log-nl> (24-05-2018).

<http://webwereld.nl/e-commerce/58525-web-log-nl-deletet-oude-blogs> (09-05-2018).

<https://web.archive.org/web/20180524115551/http://webwereld.nl/e-commerce/58525-web-log-nl-deletet-oude-blogs> (24-05-2018).

<http://webwereld.nl/e-commerce/59033-vkblogs-sneuvelen-in-nieuwste-web-log-nl-migratie> (16-05-2018).

<https://web.archive.org/web/20180524115657/http://webwereld.nl/e-commerce/59033-vkblogs-sneuvelen-in-nieuwste-web-log-nl-migratie> (24-05-2018).

<https://tweakers.net/nieuws/86711/gratis-weblogdienst-weblog-punt-nl-sluit-deuren.html> (09-05-2018).

<https://web.archive.org/web/20180524115821/https://tweakers.net/nieuws/86711/gratis-weblogdienst-weblog-punt-nl-sluit-deuren.html> (24-05-2018).

<https://www.yellowtipi.nl/blog/lowlands-voor-beginners> (16-05-2018).

<https://web.archive.org/web/20180524120049/https://www.yellowtipi.nl/blog/lowlands-voor-beginners> (24-05-2018).

<http://firstmonday.org/ojs/index.php/fm/article/view/3775> (31-05-2018). Het hele artikel staat hier

genoemd onder het adres <http://journals.uic.edu/ojs/index.php/fm/article/view/3775/3142>

<https://web.archive.org/web/20180531140657/http://journals.uic.edu/ojs/index.php/fm/article/view/3775/3142> (31-05-2018)

<https://medium.com/google-news-lab/what-is-google-trends-data-and-what-does-it-mean-b48f07342ee8> (17-05-2018).

<https://web.archive.org/web/20180524130637/https://medium.com/google-news-lab/what-is-google-trends-data-and-what-does-it-mean-b48f07342ee8> (24-05-2018).

<https://web.archive.org/web/20180528082618/http://tpo.nl/2010/12/28/verplicht-in-uw-rss-reader-Weblogs-die-er-echt-toe-doen/> (28-05-2018)

<https://www.kb.nl/bronnen-zoekwijzers/databanken-mede-gemaakt-door-de-kb/webarchief-kb/gebruiksvoorwaarden-webarchief-koninklijke-bibliotheek> (28-05-2018)

<https://web.archive.org/web/20180528133932/https://www.kb.nl/bronnen-zoekwijzers/databanken-mede-gemaakt-door-de-kb/webarchief-kb/gebruiksvoorwaarden-webarchief-koninklijke-bibliotheek>

<http://www.openschoolgemeenschapbijlmer.nl/meester-bart-is-hit-op-internet/> (04-06-2018). De oudste versie van het artikel staat op 23 augustus 2017. Er staat geen datum bij het oorspronkelijke artikel wanneer het geschreven is.

<https://web.archive.org/web/20180604122137/http://www.openschoolgemeenschapbijlmer.nl/meester-bart-is-hit-op-internet/>

<https://web.archive.org/web/20180529080216/https://marcoraaphorst.nl/51039/digitaal-erfgoed/> (29-05-2018)

Kniiff, J. van der, 'It's Alive! Eerste Nederlandse webindex gered van CD-ROM', website *Netwerk digitaal erfgoed* (18-06-2018) <http://www.netwerkdigitaalerfgoed.nl/its-alive-eerste-nederlandse-webindex-gered-van-cd-rom/>

'Selectie bij webarchivering', Website *Koninklijke Bibliotheek | Nationale Bibliotheek van Nederland* (26-11-2020) <https://www.kb.nl/organisatie/onderzoek-expertise/e-depot-duurzame-opslag/webarchivering/selectie-bij-webarchivering>

16 Artikelen gepubliceerd naar aanleiding van deze collectie

Laatst bijgewerkt op 23-02-2021

16.1 Wetenschappelijk

Geldermans, I., 2019. We blog: de casus van weblogs in webarchieven. *TMG Journal for Media History*, 22(1), pp.85–102. DOI: <http://doi.org/10.18146/tmg.438>

Iris Geldermans. (2020). De kunst van blogarchivering (Version 1). IP informatie professional, 2020(9), 27. <http://doi.org/10.5281/zenodo.4435878>

16.2 Populair

Geldermans, Iris, 'NL-blogosfeer: Burgers verslaan een ramp' website *Koninklijke Bibliotheek | Nationale Bibliotheek* (27-08-2018) <https://www.kb.nl/blogs/duurzame-toegang/nl-blogosfeer-burgers-verslaan-een-ramp>

Na aanleiding van het archiveren van de website verscheen het volgende artikel in De Stentor (Zwolle): Mollen, Janske, 'Koninklijke Bibliotheek bestempelt website oud-Zwollenaar tot "Digitaal Erfgoed"' website *de Stentor* (02-05-2018) <https://www.destentor.nl/zwolle/koninklijke-bibliotheek-bestempelt-website-oud-zwollenaar-tot-digitaal-erfgoed~acc920d/>

16.3 NL-blogosfeer wordt onderdeel van het blog

Wanneer we een blog archiveren voor de collectie kan het soms gebeuren dat deze gebeurtenis (het archiveren) onderdeel wordt van het blog (het gearchiveerde). Hieronder een lijst van posts die ik toevallig tegenkwam waar dit is gebeurd.

Website Marco Raaphorst: 'Wow! Mijn website wordt beschouwd als digitaal erfgoed! Daar word ik wel even stil van...' ⁹⁶ <https://marcoraaphorst.nl/51039/digitaal-erfgoed/>

Website Financieel Onafhankelijk Blog: 'Ik vind het dan ook bijzonder dat deze site over financieel onafhankelijk worden en blijven in Nederland in dat rijtje opgenomen gaat worden (...) Heel leuk om te zien dat de FO beweging steeds meer erkenning krijgt in Nederland!' ⁹⁷

<https://www.financieelonafhankelijkblog.nl/fob-wordt-digitaal-erfgoed-in-koninklijke-bibliotheek/>

⁹⁶ <https://web.archive.org/web/20180529080216/https://marcoraaphorst.nl/51039/digitaal-erfgoed/> (29-05-2018).

⁹⁷ <https://web.archive.org/web/20180611081057/https://www.financieelonafhankelijkblog.nl/fob-wordt-digitaal-erfgoed-in-koninklijke-bibliotheek/> (11-06-2018).

Website Stichting Beroepseer schreef naar aanleiding van selectie van hun website voor het webarchief een neutraal stuk over het beleid van de KB betreft webarchivering.

<https://www.beroepseer.nl/nl/groepsblog/algemeen/item/2606-archivering-van-website-stichting-beroepseer-door-koninklijke-bibliotheek>

17 Bijlage 1: Internetarcheologie

In deze paragraaf leg ik aan de hand van de Enschede casus uit waarom internetarcheologie belangrijk is voor deze collectie. Weblogs bevatten een schat aan archiefmateriaal wat door middel van een archiefmenu vaak goed doorzoekbaar is. Wanneer een website echter onvolledig gearchiveerd is, leidt dit tot problemen. Van de hieronder beschreven weblogs (die schreven over de ramp Enschede in het jaar 2000) is er slechts één blog wat volledig gearchiveerd is: L-rs. De andere weblogs zijn verdwenen en kunnen alleen via internetarcheologie hervonden worden. Het zijn allen historische weblogs die door Frank Meeuwsen als grondleggers worden bestempeld.

17.1 Het belang van internetarcheologie

13 mei 2000: een rustige zaterdagmiddag verandert opeens in een ramp: In Enschede vliegt een vuurwerkfabriek de lucht in: 22 mensen sterven en er zijn 950 gewonden. In *Bloghelden* schrijft Meeuwsen dat Tonie tijdens een interview opmerkt: 'De ramp in de vuurwerkfabriek vond plaats in het weekend. Toen werden de nieuwssites nog niet geüpdatet in het weekend. Heel veel redacties werkten volgens het oude media principe (...) gingen de meeste online redacties op vrijdag dicht om maandag weer te starten.' Tonie werkte zelf op dat moment voor Planet Internet. Hij was daar de eerste die weekend diensten ging draaien. Zo ook ten tijde van de vuurwerkcramp. Hij is niet de enige volgens Meeuwsen: ook Alt0169 en Esc brengen het nieuws (dat verspreid is over het internet) over de vuurwerkcramp op hun blog bij elkaar.⁹⁸

17.1.1 Tonie.net

Tonie.net is gearchiveerd in de KB, maar de archivering kwam te laat. Het archief van Tonie.net was toen al grotendeels verdwenen. In de KB Wayback Machine zitten twee snapshots van de archiefpagina van 13 – 19 mei. Beiden geven ze een melding dat het 'kladblok is dichtgeslagen' en dat de blogger is gestopt met zijn blog. In de Wayback Machine van het Internet Archive is op Tonie.net één post te vinden over de vuurwerkcramp op de dagen 13 – 14 mei. Daarnaast nog twee op 16 en 18 mei. Opvallend weinig.⁹⁹ De link van News Planet waar hij in de post op 14 mei naar verwijst, is bovendien ook door het Internet Archive te laat gearchiveerd. De pagina bestaat sinds 13 mei, maar het Internet Archive heeft een kopie gemaakt op **15 november 2001**. Het artikel is dan voor het laatst gewijzigd op 8 juni 2001.¹⁰⁰ Wat Tonie daar ooit op heeft geschreven is onmogelijk om te achterhalen.

17.1.2 ESC.

ESC is niet door de Koninklijke Bibliotheek gearchiveerd. Het Internet Archive heeft wel verschillende versies van de website.¹⁰¹ Op de homepage van 6 december kom ik een archiefknop tegen. Die archiefpagina is gearchiveerd in maart 2001. Het archief wat ik te zien krijg is echter van

⁹⁸ Meeuwsen, *Bloghelden*, 67.

⁹⁹ https://web.archive.org/web/20010830100146/http://www.tonie.net:80/index.php?p=2000_05_14_archief.html (30-05-2018).

¹⁰⁰ <https://web.archive.org/web/20011115044647/http://web.planet.nl:80/news/dossiers/vuurwerkcramp/> (30-05-2018).

¹⁰¹ <https://web.archive.org/web/20010528184124/http://esc.dds.nl:80/> (30-05-2018).

november/december van het jaar 2000. Verder terug dan 17 augustus 2000 kom ik niet. Het archief van die pagina (gearchiveerd op 30 augustus 2000) toont berichten tot 17 juni 2000. Het archief van mei 2000 is onvindbaar.

17.1.3 Alt0169

De laatste die Meeuwsen noemde is Alt0169. De KB heeft helaas alweer een latere versie van Alt gearcheveerd zonder het archief van het blog. Het Internet Archive heeft wel meerdere versies van het archief van Alt0169 (in verschillende maanden gesnapshot) waardoor ik veel berichten kan vinden over de ramp. De berichtgeving op 13 mei zelf is indrukwekkend. Op Wikipedia staat dat de eerste melding van de brand om 3 uur plaatsvindt en de explosies om half vier. Twee uur later weet Alt0169 al een bericht te plaatsen met afbeeldingen en een link naar iets wat een webcamstream lijkt te zijn. Dit bericht vult hij de rest van de avond nog aan met onder andere een ooggetuigenverslag van ene Spruijt en berichten van mainstream media zoals het AD, de Telegraaf maar ook de BBC en CNN.¹⁰² De link naar het liveblog van Spruijt is niet in het Internet Archive te vinden. De website van twente.nl is gefragmenteerd gearcheveerd. Op de gearcheveerde versie van 2 maart 2000 is namelijk de volgende home page te zien maar daarop dit bericht:

'Akkoord plan van aanpak rampwijk Eerste gesprek van de Belangenvereniging Slachtoffers Vuurwerkramp met B&W; van Enschede leidt tot plan van aanpak voor de wederopbouw van verwoeste wijk. Akkoord bereikt over inspraak bewoners, herstel, herbouw en terugkeer naar eigen woningen.'¹⁰³

Dit kan niet van 2 maart 2000 zijn. De vuurwerkramp vond plaats op 13 mei 2000.

Naast de reguliere nieuwsberichtjes quote Alt ook apartere berichten zoals berichten van Grolsch, waarvan de fabrieken ook in brand waren gevlogen.¹⁰⁴ Omdat Alt goed gearcheveerd is door het Internet Archive kun je ook vinden wat er in de dagen daarna door Alt allemaal geplaatst wordt. Het is echter wel lastig navigeren omdat de website zo gebroken gearcheveerd is.¹⁰⁵

Deze drie bovengenoemde websites zijn op het moment niet te archiveren omdat zij niet langer op internet staan. Er is wel wat in het Internet Archive te lezen, maar moet je erg oppassen dat je niet op het verkeerde been gezet wordt. Bovendien moet je geluk hebben dat datgene wat je precies zoekt gearcheveerd is. De enige manier om de zo compleet mogelijke informatie van een offline websites te achterhalen is via internet archeologie. Internet archeologie kan ervoor zorgen dat de archieven van deze websites herontdekt kunnen worden zodat we kunnen zien wat Tonie, Esc en Planet Internet op de dag zelf schreven over de vuurwerkramp in Enschede. Deze primaire bronnen plaatsen bovendien de gebeurtenis in context van de tijd zelf omdat bloggers zich continue bezighouden met de

¹⁰² <https://web.archive.org/web/20010421110602/http://www.alt0169.com:80/0051.shtml> (30-05-2018).

¹⁰³ <https://web.archive.org/web/20000302143536/http://www.twente.nl:80/> (30-05-2018).

¹⁰⁴ <https://web.archive.org/web/20010421110602/http://www.alt0169.com:80/0051.shtml> (30-05-2018). De berichten zijn van 14 en 15 mei.

¹⁰⁵ <https://web.archive.org/web/20010301201159/http://www.alt0169.com:80/0052.shtml> en <https://web.archive.org/web/20010421110602/http://www.alt0169.com:80/0051.shtml> (18-06-2018).

actualiteit. Zo zit er tussen alle vuurwerkrap posts van Alt0169 ook opeens een opmerking over de Gladiator film die net is uitgekomen.

17.1.4 L-rs

Deze website is een voorbeeld van wat (nog) wel te archiveren is. De website van L-rs bevat een archief waar ook mei 2000 nog inzit. Hierin is bijvoorbeeld zijn verslaggeving van de gebeurtenissen op 14 mei nog terug te lezen. Van de verwijzingen naar Enschede nieuws doet geen enkele het meer. Het archief van L-rs is op het moment nog wel online ondanks dat de website al jaren niet meer actief is. Zijn blik op (onder andere) de gebeurtenissen in Enschede zal in de toekomst terug te lezen zijn in deze collectie.¹⁰⁶ Het feit blijft wel dat dit geen precieze weergave is van de website op zondag 14 mei 2000. Wij archiveren immers het archief van de L-rs website uit 2010. De balk aan de zijkant van de website kan door de tijd heen veranderd zijn. Misschien is ook de opmaak van de website door de tijd heen veranderd. Of het bericht tussendoor nog bewerkt is, weten wij ook niet.

17.2 Internetarcheologie wishlist

Een lijst van websites die niet langer online zijn maar wel belangrijk zijn voor het corpus. Deze zouden via internet archeologie gevonden en gereconstrueerd kunnen worden.

- Loglands Eerste moblog / liveblog in Nederland (gemaakt tijdens Lowlands 2004)
- Frank-ly.nl Komt in de literatuur regelmatig naar voren
- Daily Wacko Volgens Wikipedia de eerste blog van Nederland en een parodie op de Daily Planet.
- About:blank

Eerste blogtijdschrift in Nederland (2002 opgericht) met veel bekende bloggers die een bijdrage leveren. Groeit om tot een weblog (vanaf eind 2003). Het is bovendien vanaf 2003 de organiserende partij achter de Dutch Bloggies. About:blank gaat in 2007 op in de stichting Dutch Bloggies.¹⁰⁷

Grondleggers van het bloggen (volgens Frank Meeuwssen in *Bloghelden*) waarvan de websites op dit moment nog onvindbaar zijn¹⁰⁸:

- E-Wave
- Daily Planet
- Het Dagelijks Nieuws
- Smoel.nl
- Pjoe.net
- sikkema.net
- VanTienen.com
- Weblogger.nl
- ISDIB
- Hagenaars.net
- Linked
- TheWolf

¹⁰⁶ http://l-rs.org/archief/2000_05.html (04-06-2018).

¹⁰⁷ Meeuwssen, *Bloghelden*, 147 – 151.

¹⁰⁸ Meeuwssen, *Bloghelden*, 47.

- MoreWhereThat
- Joglog
- :]=Esc
- MisterEdBlog
- Stronk
- Cyberramon
- Zo.nu/&.dan
- Planeet Eendvogel

Grondleggers van het bloggen volgens Frank Meeuwsen in *Bloghelden* die op dit moment wel door de KB gearchiveerd zijn, maar onvolledig zijn / archieven missen ¹⁰⁹:

- Tonie.net (& Planet Internet waarvoor hij schreef)
- Alt0169

Ondanks dat Alt0169 het beste in het Internet Archive zit, zou ik deze website ook in het archief van de KB willen hebben. Ten eerste omdat het dan zeker is dat de website duurzaam bewaard wordt voor de toekomst. Ten tweede zou het een goed controlemiddel zijn om te vergelijken wat het Internet Archive archiveert tegenover een reconstructie via internet archeologie. Hierdoor kunnen eventuele gaten (in het IA en de archeologiemethode) en missende / misleidende informatie worden ontdekt.

¹⁰⁹ Meeuwsen, *Bloghelden*, 47.

18 Bijlage 2: Definitie uit gelezen literatuur

Deze tabel bestaat uit een aantal delen van definities die gevonden zijn in de volgende zes bronnen:

- Bekker, C., *De blog bijbel. Voor beauty chicks & shopaholics* (Maartensdijk 2014) 9 – 10, 15, 24 – 25, 27
- Deriemaeker, K., *Blog boek. Tips & inspiratie voor meer blog- en vlogplezier* (Antwerpen 2017) 24 en 259.
- Duval, S., *How blogs work. Tips & tricks voor beginnende en ervaren bloggers* (Tielt 2014) 9, 13 - 14.
- Holthuijsen, D., *Weblogs. Uw persoonlijke blog op het web!* (Haarlem 2005) 10 en 36.
- Meindersma, C., *Wetboek voor bloggers. De (on)geschreven regels voor bloggers, journalisten en content creators* (Den Haag 2016) 25, 28 - 29.
- <https://nl.wikipedia.org/wiki/Weblog> geraadpleegd op 9 april 2018.

Legenda:

X gebruikt dit element om de term 'blog' te definiëren

- Benoemd dit element niet expliciet

Tekst Gebruikt een andere (soortgelijke / nadrukkelijke / genuanceerde) term om dit te beschrijven

Grijze kleur: Element is opgenomen bij de paragraaf 'definitie blog' als veel voorkomend element (maar geen verplicht onderdeel)

Groene kleur: Element is opgenomen bij de paragraaf 'definitie blog' als vast onderdeel.

Zwarte kleur: Element is niet opgenomen in de paragraaf 'definitie blog'.

	Wikipedia	Holthuijsen	Duval	Bekker	blogboek	wetboek
Reeks afzonderlijke berichten (posts)	-	X	X	-	X	-
Korte stukjes	Microblog?	-	X	X	-	-
Tijd en datum staan erbij genoemd	-	X	-	-	-	-
Meest recente post staat bovenaan	X	X	-	-	X	X
Regelmatig bijgewerkt	X	X	X	X	X	X

	Wikipedia	Holthuijsen	Duval	Bekker	blogboek	wetboek
Dagelijks bijgewerkt	X	X	X	liefst	-	-
Wekelijks	-	-	X	-	-	-
Bezoekers kunnen reageren op het blog	meestal	X	-	X	X	X
Bevat teksten, afbeeldingen en/of geluid	Meestal tekst maar ook...	X	-	-	-	-
Bevat tekst, foto's en filmpjes	Meestal tekst maar ook...	-	X	-	-	-
Bevat tekst, beelden en links	-	-	-	-	X	X externe links
Heeft een archief waarin alle bijdragen bewaard worden	-	X	-	-	-	-
Heeft koppelingen naar andere websites of blogs	-	X	-	-	X	-
Eigen domeinnaam of onderdeel van een weblog service	-	X	-	-	-	-
Marketinginstrument	-	X	-	-	-	-
Journalistiek middel	-	X	-	-	-	-
Communicatie middel	-	X	-	-	-	-
Politiek middel	-	X	-	-	-	-
Levert achtergrond informatie over hobby of andere bezigheden	specialisatie	X	-	-	-	-
Actualiteit staat centraal	-	X	X	-	-	X
Blog is absoluut persoonlijk, in tegenstelling tot algemene nieuwssites.	Persoonlijk dagboek	-	X	Schrijf persoonlijk!	XXXX	X
Website	X	X	X	-	X	X /Online
Deel van een website	-	-	-	-	-	X
Is subjectief	-	-	X	-	Mening bloggers Centraal	-

	Wikipedia	Holthuijsen	Duval	Bekker	blogboek	wetboek
Geen twee blogs zijn hetzelfde	-	-	X	-	-	-
Qua thema: alles is mogelijk	-	-	X	-	-	-
Publiek kan van over de hele wereld komen	-	-	X	-	-	-
Vast publiek	X	-	-	-	-	-
Doelgroep kan veranderen	-	-	-	X	-	-
Bloggers lezen veel blogs waardoor nieuws zich snel verspreid op dit medium	X	-	-	-	-	-
Blog is openbaar	-	-	-	X	-	-
Kan een fulltime baan zijn	-	-	-	X	-	-
Iedereen kan bloggen	-	-	-	X	-	-
Instant verspreiding door op 'publiceer' te drukken	-	-	-	-	X	-
'Een blog is een gids die de lezer mee op sleeptouw neemt om de lezer dingen te tonen die hij voor jou heeft uitgekozen.' ¹¹⁰	-	-	-	-	X	-
Persoonlijk, zakelijk, opiniërend of objectief	-	-	-	-	-	X

¹¹⁰ Deriemaeker, K., *Blog boek. Tips & inspiratie voor meer blog- en vlogplezier* (Antwerpen 2017) 24.

19 Bijlage 3: Thema's / genres: welke categorieën houden anderen aan?

Bronnen gebruikt:

- Holthuijsen, D., *Weblogs. Uw persoonlijke blog op het web!* (Haarlem 2005) 18 - 35.
- Duval, S., *How blogs work. Tips & tricks voor beginnende en ervaren bloggers* (Tielt 2014) 14.
- <https://nl.wikipedia.org/wiki/Weblog> geraadpleegd op 9 april 2018
- Bekker, C., *De blog bijbel. Voor beauty chicks & shopaholics* (Maartensdijk 2014) 11 – 12.
- Deriemaeker, K., *Blog boek. Tips & inspiratie voor meer blog- en vlogplezier* (Antwerpen 2017) 25 – 27.
- Meindersma, C., *Wetboek voor bloggers. De (on)geschreven regels voor bloggers, journalisten en content creators* (Den Haag 2016) 31 – 34.

Legenda:

X benoemd dit thema expliciet

- Benoemd dit element niet expliciet

Tekst Gebruikt een andere (soortgelijke / nadrukkelijke / genuanceerde / overkoepelende) term om dit te beschrijven

1 / 2 / 3 Er worden in de bron meerdere indelingen gehanteerd: term wordt gebruikt in indeling 1, indeling 2 of indeling 3.

Bij het *Blog boek* is indeling 1 op basis van technieken, indeling 2 op basis van doel/inhoud en noemen zij ten derde nog een aantal kleinere blogs.¹¹¹

Bij het *Wetboek voor bloggers* is indeling 1 op basis van mediatype, indeling 2 op basis van doelstelling en indeling 3 op basis van genre. Ook hier worden daarnaast nog veel meer kleinere blogs benoemd.¹¹²

	Wikipedia	Holthuijsen	Duval	Bekker	blogboek	wetboek
Babysite	X	-	-	-	-	-
Mamablog	X	-	-	-	3	
Beauty	-	-	Lifestyle	X	3	

¹¹¹ Meindersma, C., *Wetboek voor bloggers. De (on)geschreven regels voor bloggers, journalisten en content creators* (Den Haag 2016) 31 – 34.

¹¹² Deriemaeker, K., *Blog boek. Tips & inspiratie voor meer blog- en vlogplezier* (Antwerpen 2017) 25 – 27.

	Wikipedia	Holthuisen	Duval	Bekker	blogboek	wetboek
Blawgs (rechtspraak)	-	-	-	-	3	X
Blog over blogs	-	-	-	-	3	
Boekenblogs	-	-	-	-	3	
Clublog	X	-	-	-	-	
Cultuurlog (en kunst)	-	X	Lifestyle	-	-	
Commercieel blog	-	-	-	-	-	2
Corporate blog of Businessblog	X	X	-	-	Bedrijfsblog 3	2 zakelijk
Dierenblog	-	-	-	-	3 kat/hond	
Design	-	-	Lifestyle	-	3	
DIY blog	-	-	-	-	3	
Fashionblog	X	-	Lifestyle	Modeblog	3	
Filmblog	X	-	-	-	-	
Foodblog	X	-	Lifestyle	-	3	
Fotoblog	X	X	Lifestyle	-	1 photolog	1
Gameblogs	-	-	-	-	3	
Humorlogs	-	X	-	-	3 mopjes	
Klasblog	X	Educatiefblogs	-	-	3	2 edublogs
Lifelog	X	X	-	Persoonlijke diaryblog	Persoonlijk blog 3	2persoonlijk blog
Linkdump of linklog	X	X	-	-	1	
Microblog	Facelog / space-blog	-	-	-	1	1
Moblog	X	X	-	-	-	
Muzieklog	X	X (en audio)	-	-	-	
Nieuwslog	X	Journalistiek blog	-	-	-	2
Podcast	-	-	-	-	-	1
Politiek weblog	X	X	-	-	3	
Programmeerblogs	-	-	-	-	3	
Reisblogs	X	Vakantielogs	Lifestyle	-	3 travel	

	Wikipedia	Holthuijsen	Duval	Bekker	blogboek	wetboek
Schrijfblogs	-	-	-	-	3	
Shocklog	X	X	-	-	-	
Seniorenblogs	-	-	-	-	3	
Sportlogs	-	X	-	-	3 loopblog voetbal	
Stadsblog	X	-	-	-	3 cityblog	
Superblogger	X	-	-	-	-	
Tech blog	-	-	-	-	3	
Tekstblog	-	-	-	-	-	1
Themablog	X	X	-	-	-	
Tuinlogs	-	X	-	-	3 moestuin	
Tv-blog	X	-	-	-	-	
Vrijheid meningsuiting	-	X	-	-	-	
Vlog	wijkvlog	-	-	-	1	1
Wetenschapslog	-	X	-	-	-	

Extra opmerkingen over de bronnen:

Duval: behandelt in haar boek alleen de categorie lifestyle blogs, waar allerlei soorten thema's in te onderscheiden zijn.

Bekker: behandelt in haar boek slechts een beperkt genre aan blogs. Op de cover staat echter dat een blog in principe over alles kan gaan. Ook politiek.

Wetboek: onderscheidt ook nog andere soorten blogs dan op genre of format:

- Nicheblogs: over één / enkele onderwerpen Omniblogs: over een heleboel onderwerpen
- Individuele blog Groepsblog (Multi Author Blogs)
- Gastblogs: blogs die door iemand anders geschreven zijn dan door de persoon of personen die normaal de content verzorgen

Indeling Martijn Kleppe voor zijn onderzoek:

Deze categorieën komen voort uit het onderzoek: 'Analysing & understanding news consumption patterns by tracking online user behaviour with a multimodal research design' geschreven door Martijn Kleppe en Marco Otte.¹¹³ Deze indeling was echter wel bedoeld voor websites in zijn algemeen en niet specifiek voor blogs.

auto's	educatie & opleiding	kinderen	sport
banken & beleggen	elektronische webdata	koken	telecom
bedrijven	email	media	vakantie
bloggen	energie	naslagwerk	verenigingen
carrière	games	nieuws	vergelijking
communities	gezondheid	overheid	verzekeraars
consumentenelektronica	goede doelen	religie	video, muziek en radio
cultuur	gokken	shopping	vrije tijd
daten	huisvesting	social media	weer & verkeer
			Zoekmachine

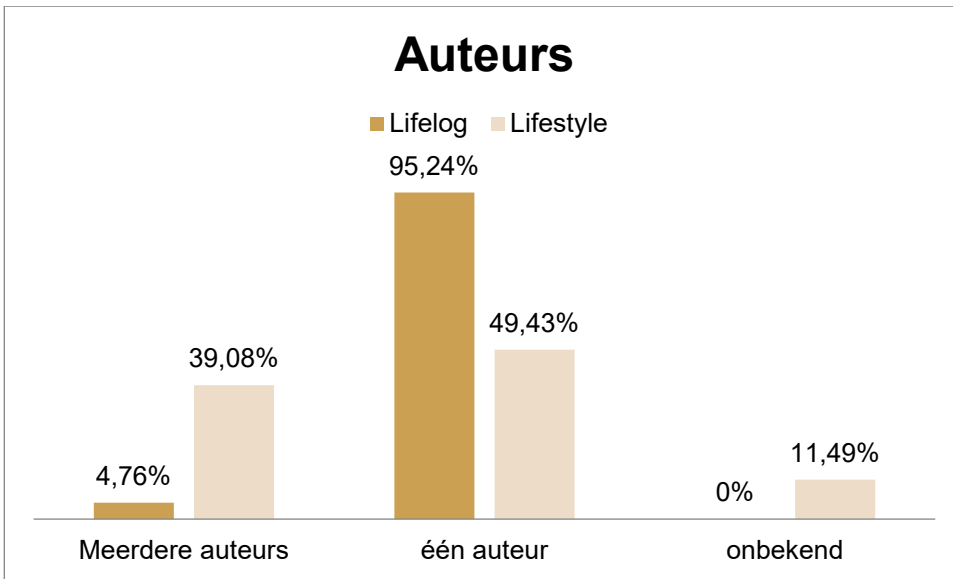
¹¹³ Martijn Kleppe, Marco Otte, Analysing and understanding news consumption patterns by tracking online user behaviour with a multimodal research design, *Digital Scholarship in the Humanities*, Volume 32, Issue suppl_2, December 2017, Pages ii158–ii170, <https://doi.org/10.1093/llc/fqx030>.

20 Bijlage 4: Lifestyle VS Lifelog analyse

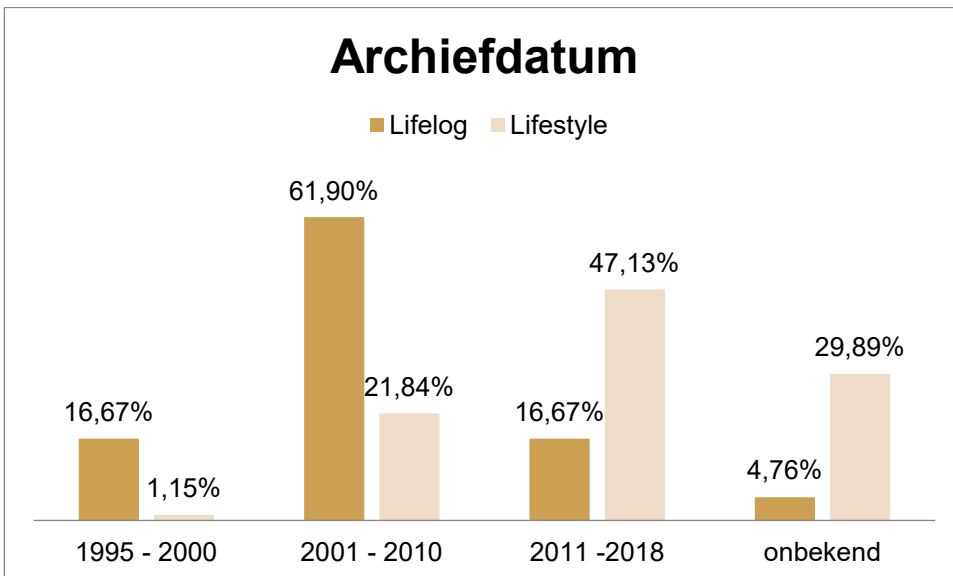
Analyses in deze collectiebeschrijving zijn gemaakt op basis van de selectielijst van 20-06-2018. Deze selectielijst wijkt af van de huidige NL-blogosfeer collectielijst.

Vergelijking Lifestyle (87 records) en Lifelog (42 records) in percentages:

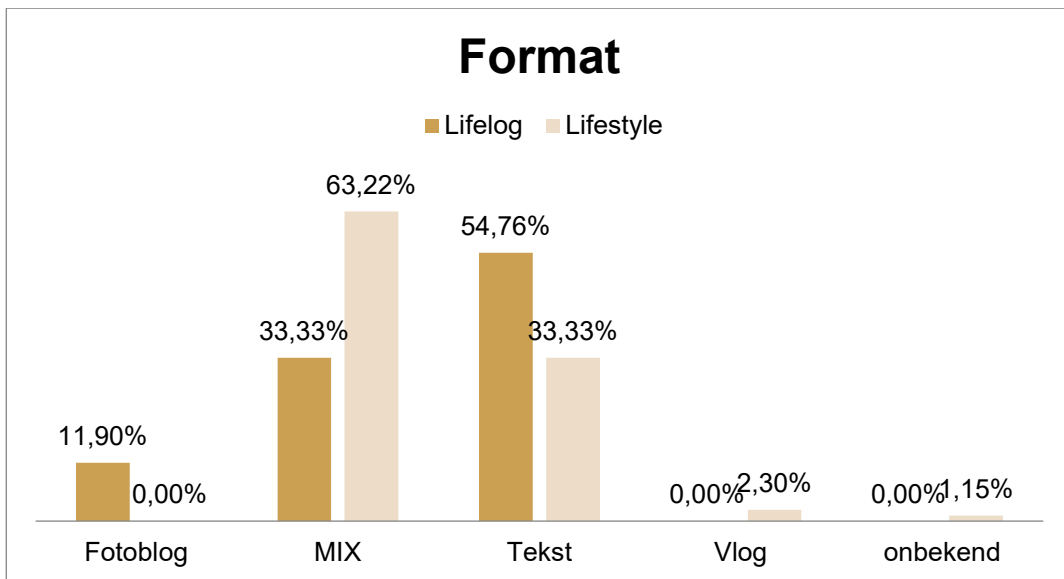
Auteurs	Lifelog	Lifestyle
Meerdere auteurs	4,76%	39,08%
Eén auteur	95,24%	49,43%
Onbekend	0%	11,49%
Format	Lifelog	Lifestyle
Fotoblog	11,90%	0,00%
MIX	33,33%	63,22%
Tekst	54,76%	33,33%
Vlog	0,00%	2,30%
Onbekend	0,00%	1,15%
Archief	Lifelog	Lifestyle
1995 - 2000	16,67%	1,15%
2001 - 2010	61,90%	21,84%
2011 -2018	16,67%	47,13%
Onbekend	4,76%	29,89%



Lifestyle wordt vaker door één auteur geschreven, maar het verschil tussen één en meerdere auteurs is veel kleiner dan bij Lifelog: bij Lifelogs wordt het blog bijna altijd door één persoon geschreven.



Lifelogs komen meer voor in 1995 – 2000 dan Lifestyle (zoals verwacht). Historisch gezien waren blogs dagboeken van de surftochten. Lifestyle komt op in 2001 – 2010 waarbij het piekt in 2011 – 2018. (in deze collectie).



Lifestyle blogs **binnen de collectie** zijn vooral een mix van elementen: tekst, filmpjes en foto's. Lifelogs blijken toch voornamelijk uit tekst te bestaan.

Absolute aantallen:

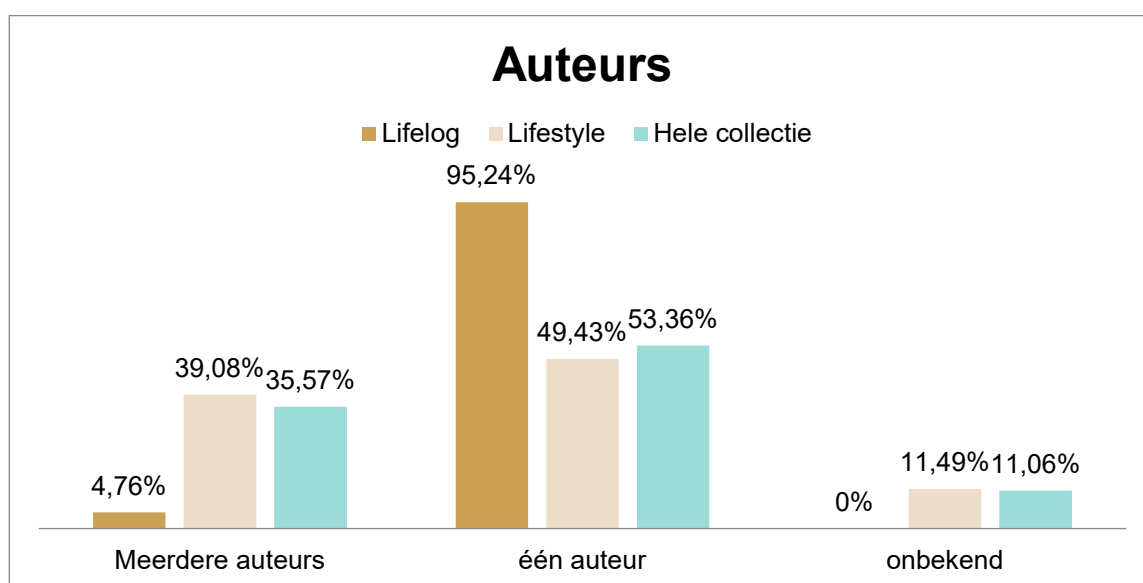
Lifelog:

Auteurs	Lifelog	In percentage
JA	2	4,76%
NEE	40	95,24%
Format	Lifelog	In percentage
Fotoblog	5	11,90%
MIX	14	33,33%
Tekst	23	54,76%
Archief	Lifelog	In percentage
1995 - 2000	7	16,67%
2001 - 2010	26	61,90%
2011 -2018	7	16,67%
Onbekend	2	4,76%

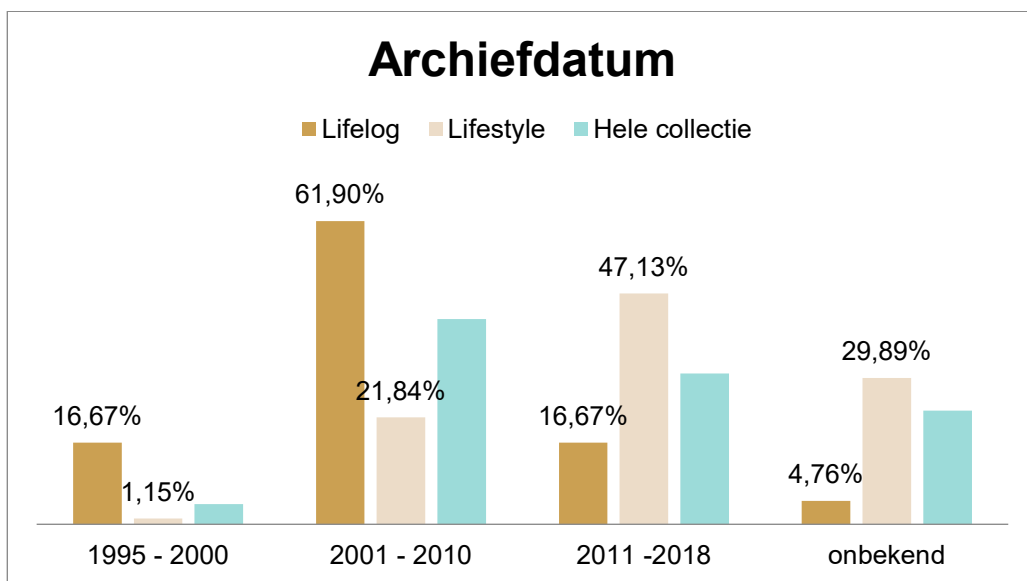
Lifestyle:

Meerdere auteurs	Lifestyle	In percentage
JA	34	39,08%
NEE	43	49,43%
Onbekend	10	11,49%
Format	Lifestyle	In percentage
MIX	55	63,22%
Onbekend	1	1,15%
Tekst	29	33,33%
Vlog	2	2,30%
Archiefdatum	Lifestyle	In percentage
1995 - 2000	1	1,15%
2001 - 2010	19	21,84%
2011 - 2018	41	47,13%
Onbekend	26	29,89%

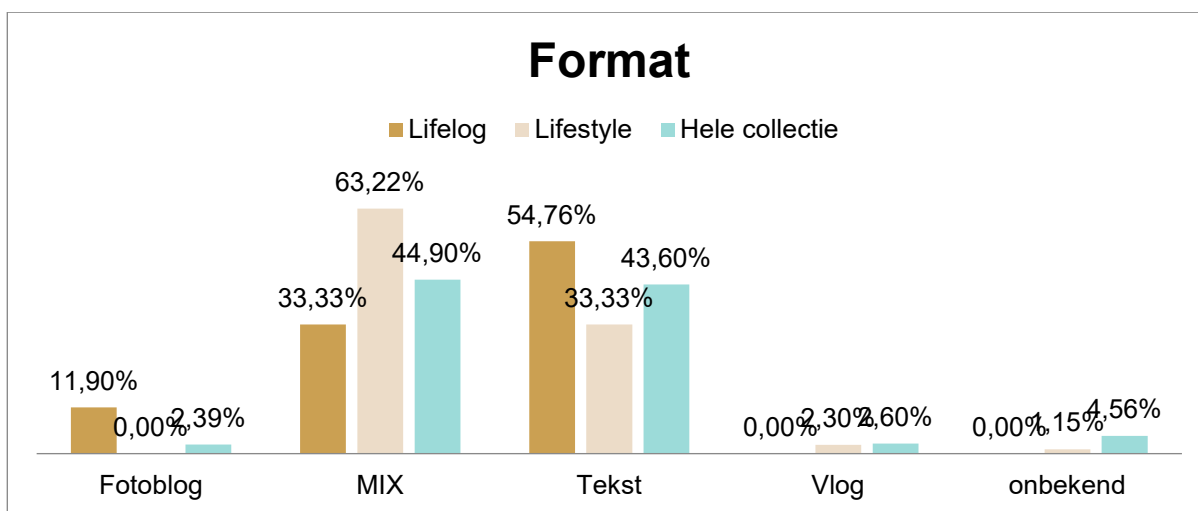
Analyse met het gemiddelde percentage van de hele collectie erbij:



Lifelog is hier opvallend vaak boven of onder het gemiddelde.



Lifestyle en Lifelog liggen hier opvallend vaak boven en onder het gemiddelde.



Lifestyle is opvallend vaak MIX

Lifelog is meer tekst, maar ook opvallend vaak een fotoblog.

