
Ronde 8

Maarten Sprenger & Carsten Schnober
Slim Zoeken Amsterdam
Contact: maarten@slimzoeken.nu
carschno@gmail.com

Zoeken onderzocht – Het 100 Queries project

1. Inleiding

Terwijl de wereld overspoeld wordt met SEO-tactieken¹, wordt er vrijwel geen aandacht besteed aan het perspectief van de ontvanger. Er bestaat nauwelijks onderzoek

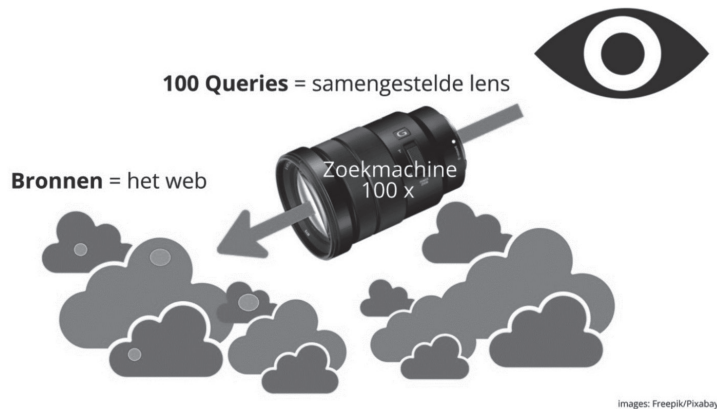
naar wat online beschikbaar is voor de informatieve consument en naar wat daarvan gepresenteerd wordt door zoekmachines. Er bestaat evenmin recent onderzoek naar hoe je, als leerling, zoekmachines het best kunt gebruiken².

Aan dit beeld verandert de komst van chatbots niets. Zij maken immers gebruik van het gehele web plus aanvullingen. Sterker nog: het wordt feitelijk steeds belangrijker om de uitkomsten van chatbots, die niet bekend staan om hun inherente betrouwbaarheid, te kunnen controleren. In het onderwijs moet dus vaker gesproken gaan worden over de kwaliteit van het web, het aanwijzen van vertrouwde bronnen en over een gericht gebruik van zoekmachines.

In 2023 deden Sprenger & Schnober onderzoek naar de Google-zoekresultaten op 100 zoekopdrachten van basisschoolleerlingen. We wilden hiermee:

1. een beeld krijgen van de beschikbare en via Google aangeboden informatie op het web;
2. een model opzetten om de kwaliteit van getoonde internetinformatie bespreekbaar te maken en te beoordelen.

100 authentieke *queries* zijn onze lens op wat er op het web beschikbaar is en wat er via zoekmachines wordt gepresenteerd aan de zoeker (in de vorm van 1.000 resultaten).



Figuur 1 – 100 Queries.

2. Dataverzameling

We kozen 100 uiteenlopende kinderzoekopdrachten (*'queries'*) uit een random set van 200³. We legden databases aan voor *queries*, zoekresultaten en bronnen. De veldwaarden van de *query*-database waren vrij exact: lengte, soort *query*, leeftijd, onderwerp.

De velden voor zoekresultaten en bronnen voor onder meer commercialiteit, informatiesoort, toegankelijkheid en betrouwbaarheid moesten speciaal voor dit onderzoek geformuleerd en ingesteld worden, zowel de veldnamen als de waarden.

We analyseerden de eerste tien resultaten onder de tab 'Alle'. We maakten onderscheid tussen 'de inhoud', gerepresenteerd door de webpagina's die zoekmachines weergeven als zoekresultaten, en 'de bronnen', waar deze pagina's deel van uitmaken.

Zoekresultaten

- Relevantie: past het resultaat bij de zoekopdracht?
- Ranking van de zoekresultaten van 1-10.
- Toegankelijkheid (combinatie van leesbaarheid en webontwerp; expert-rating).

Bronnen

- Aanwezigheid en kwaliteit van 'over-ons-/about-pagina's'.
- Doel van de bron.
- Sector.
- Commercialiteit.
- Hoeveelheid reclame.
- *Blocked trackers*.

3. Wat opvalt

- Wiki's zijn graag geziene bronnen in Google. Wikipedia in de eerste plaats (10% van de resultaten, een derde daarvan staat bovenaan), maar ook Wikikids (2,4%, maar nooit bovenaan) en Wiktionary (1%). In het geval van Wikikids is dat dubieus te noemen: het volume van Wikikids is wel groot (45.000+ pagina's), maar de kwaliteit is onvoldoende.
- 7,7% van de zoekresultaten is een Google-verwijzing naar een andere tab (zoals video, afbeeldingen of plaatsen).
- 40% van de bronnen is *not-for-profit*.
- Daaronder zitten veel overheidsbronnen (NL + BE): 75 resultaten (7,5% van 998) van 55 bronnen (10% van 542 bronnen) voor onderwerpen als water, toerisme, gezondheidsvoorlichting en regelgeving.
- De toegankelijkheid van de overheidsbronnen gaat de goede kant op (een 50/50-verhouding van redelijk begrijpelijk versus complex).
- Veel bronnen hebben goede 'about-pagina's' (77%); er is ook een duidelijke relatie tussen dubieuze bronnen en het ontbreken van zo'n 'over-ons pagina'.
- 5% van de bronnen is ultra-commercieel (*aggregators*, *clickbait*, te veel *trackers*).

4. De commercialiteit

Het is niet eenvoudig te zeggen wat goed en slecht is in het waarderen van online informatie. We hebben een behoorlijk gedetailleerde metadatering opgezet, maar buiten onder andere Wikipedia, Schooltv en Thuisarts zijn er talloze afleidende, misleidende en privacy-ondermijnende prikkels ingebakken in het gebruik van online informatie. Veel websites zou je überhaupt niet willen tonen aan een basisschoolleerling. En dat heeft maar bij één zoekopdracht met porno te maken. Veel bronnen achter de zoekresultaten dienen allereerst een commercieel doel: puur webshop, bedrijfsinformatie en zelfpromotie. Of de bronnen hebben advertentie-inkomsten als hoofddoel en informatie geven als bijzaak. Die zelfpromotie komt bij de helft van de bronnen voor, o.a. met *calls to action* zoals *pop-ups* voor nieuwsbrieven en aanbiedingen. Je zou kunnen zeggen dat je daaraan moet wennen, dat het bij het web hoort. Maar het is afleidend. Een leerling heeft er helemaal niks aan. Het is storende ruis.

Commercie en promotie tonen zich als volgt (in verschillende combinaties):

1. Bedrijven en instellingen die vooral iets verkopen of hun activiteiten beschrijven: 60% van de getoonde bronnen (slechts zeer ten dele als antwoord op de mogelijke zoekintentie).
2. Advertenties op websites: 24% van de getoonde bronnen.
3. Promotie doelen: informatie publiceren om bezoek te genereren: 25% van de getoonde bronnen.
4. Bedrijfsuithangbord (“dit zijn onze producten/diensten”; “hier kun je ons vinden”). Op zich een reëel doel voor een bedrijf, maar als informatief resultaat zelden relevant voor een kinderquery: 21%.
5. *Company-promotion*: bezoekers laten weten wat je als bedrijf nog meer in huis hebt): 54%.
6. (*Tracking*) cookies plaatsen en profielen opbouwen van bezoekers: slechts 4% heeft geen *geblockte trackers*; 96% dus wel!

Dit resulteert in maximaal 30% kans op een eenvoudige informatie-ervaring voor een leerling.

5. Zoeken voor leerlingen

De gevonden bruikbare kinderinformatie is minimaal of zit enigszins verstopt. Samenvatting rond bronnen gemaakt voor 8-14 jaar:

- In totaal is 5% van de bronnen een po/vo bron (29 van de 542). Daarvan zijn er 16 gemaakt voor het po, incl. 5 onbetrouwbare bronnen. Conclusie: 2% van de 542 gevonden bronnen is direct goed te gebruiken door basisschoolleerlingen.

- Schooltv, Jeugdjournaal, Klokhuis e.a. zijn helaas alleen te vinden onder Google's video-tab, voor toch gelukkig 38% van de *queries*. Echter: vooral via hun YouTube-kanalen, niet met hun grotere eigen volume. En Jeugdjournaal en Klokhuis zijn nooit te zien onder de algemene tab 'Alle'.
- Docukit komt 4x voor als resultaat, maar deze bron is al heel lang niet geüpdatet.
- Wikikids komt 24x voor (nooit bovenaan), maar is geen geschikte bron.
- *Dierenwiki* komt 5x voor, maar is een nepbron (*aggregator*).

6. Aanbevelingen voor het onderwijs

Het is essentieel dat leerlingen vanaf een jaar of tien zicht krijgen op de uiteenlopende bronnen zoals we die tegenkomen via een zoekmachine, en op de motieven van de makers. Gelukkig is dat doel ook expliciet in de 'Actualisatie Kerndoelen' (SLO 2023) opgenomen (zie: 'Digitale geletterdheid, p. 26: A. Praktische kennis & vaardigheden Kerndoel 2. Digitale media en informatie'; 'Nederlands, p. 15; A. Communicatie, A1. Taalvaardigheid Kerndoel 7').

Aan de hand van de uitkomsten van ons onderzoek kunnen keuzes gemaakt worden voor de opzet van gesprekken over het gebruik van vertrouwde bronnen en van zoekmachines.

Als eerste zouden we willen voorstellen dat elke docent die een (nieuwe) online bron laat zien in de klas (voor welk doel dan ook), er steeds bij vertelt wie de makers van de informatie zijn en wat het doel van de bron is. Dit, als goede gewoonte om leerlingen ervan te doordringen om dat ook altijd zelf uit te zoeken. En zeker ook om ervaring op te doen met het onderzoeken daarvan, want het web is schimmig. De velden en labels uit ons onderzoek kunnen daarbij behulpzaam zijn¹.

Het is ook essentieel om samen een vocabulaire te ontwikkelen voor het bespreken van webcontent in het algemeen. En zeker voor het web is het hoog tijd om te stoppen met het idee dat alle basisschoolleerlingen pre-academische kwaliteiten als bronnen vergelijken zouden kunnen ontwikkelen. Wijs leerlingen vooral de weg. En laat leerlingen niet zonder voorbereiding googelen of chatten met chatbots.

7. Samenvatting voor het onderwijs

1. *Maak een zoekopdracht specifiek, gebruik meer woorden*
 - Zoeken met één woord is weinig succesvol: je krijgt dan veel webshops en andere *transactional* resultaten.
 - Kindercontent komt niet erg naar boven; herformuleren is dus belangrijk.
 - Zoek via de video-tab naar goede kindervideo's. Ga ook voor Schooltv naar de video-tab of typ bijvoorbeeld "koraalrif schooltv".

2. *Start rechtstreeks bij Wikipedia of bij een vertrouwde kinderbron*
 - Webkennis is essentieel. Zomaar googelen is vaak tijdverspilling en misleidend.
 - Leer vertrouwde bronnen ontdekken/kennen.
3. *Zoek altijd naar de makers-pagina voor nieuwe/onbekende bronnen*
 - Veel bronnen hebben prima ‘about-pagina’s’. Zoek daar naar de maker(s) en het doel van een bron.

Tot slot wordt het gebruik van een *tracking cookies blocker* zoals uBlock Origin warm aanbevolen.

Referenties

- Allison Druin, A. & E. Foss (2014). *Children's Internet Search*. San Rafael, California: Morgan & Claypool Publishers.
- Jochmann-Mannak, H. (2014). *Websites for children: Search strategies and interface design*. Twent: UT Twente. [Proefschrift].
- SLO (2023). ‘Actualisatie Kerndoelen’. Online raadpleegbaar op: www.actualisatiekerndoelen.nl; www.actualisatiekerndoelen.nl/digitalegeletterdheid; www.actualisatiekerndoelen.nl/nederlands.
- Sprenger, M. (2018). *Slim Zoeken op internet*. Groningen: Uitgeverij De Wereld Van Bovenaf / Slimzoeken.nu.

Noten

- ¹ SEO staat voor ‘*Search Engine Optimization*’. SEO zorgt ervoor dat je als website op een hoge plek in een algemene zoekmachine komt.
- ² Het inspirerende onderzoek *Children's Internet Search* van Druin & Foss dateert uit 2014 en heeft vrijwel geen vervolg gekregen.
- ³ De sample set dateert uit 2022 en werd ons ter beschikking gesteld door Thijs Westerveld van WizeNoze.
- ⁴ Het volledige onderzoeksverslag is te vinden op GitHub: <https://github.com/SlimZoeken/100queries>.