

Onderzoeksverslag Martijn Bankert 2019

Inhoud

Introductie	3
SEO	3
SPA's	3
Het probleem	4
Onderzoeksvraag	5
Deelvragen	5
Hoe wordt een website gerenderd?	6
Wat doet daarin de rendering engine?	6
Hoe wordt een SPA gerenderd?	7
Client-side rendering (CSR)	7
Server-side rendering (SSR)	8
Prerendering	8
Even vergelijken	9
Wat is Google crawling, indexing en SEO-ranking?	10
Google crawling	11
Google indexing	12
Google SEO-ranking	12
Waar wordt de SEO-score op gebaseerd?	13
Technical (Hoge prioriteit)	13
Content (Medium prioriteit)	13
User Experience (Lage prioriteit)	14
Mobile (Medium prioriteit)	14
Welke SEO-technieken kunnen we anders of beter toepassen voor SPA's?	15
Conclusie	16
Bronnenlijst	17
Wetenschappelijke bronnen	17
Overige bronnen	17

Introductie

Om bij het begin te beginnen, wat is eigenlijk een website? Onder een website verstaan we een verzameling samenhangende webpagina's met gegevens zoals: afbeeldingen, teksten en video's [4]. Deze verzameling samenhangende webpagina's wordt vervolgens opgeslagen of gehost¹. Dit soort websites worden ook wel MPA's, Multi Page Applications, genoemd.

SEO

Nu weten we wat een website is, maar wat is dan SEO? SEO staat voor Search Engine Optimization of in het Nederlands zoekmachineoptimalisatie. SEO is het actief optimaliseren van een website om het bezoeken van jouw website of app, op relevante zoekwoorden vanuit zoekmachines zoals Bing, Google, Yahoo (etc.), te verhogen [6]. Heel kort gezegd is SEO nodig voor de vindbaarheid van je website voor zoekmachines. Hoe beter je je SEO geregeld hebt, hoe hoger je website scoort op vindbaarheid en daardoor kom je hoger in de ranglijsten te staan van de zoekmachines.

SPA's

Als laatste hebben we nog de Single Page Applications (SPA's), SPA's zijn websites die gemaakt worden met bijvoorbeeld: Angular, React of Vue. Deze websites/applicaties maken gebruik van een principe waarin je een webpagina dynamisch kan herschrijven. Zodra je op een MPA ergens op klikt wordt de gehele pagina opnieuw geladen vanaf de server. Met een SPA wordt alleen bepaalde content opnieuw ingeladen, daardoor heeft de gebruiker geen last van onderbrekingen tijdens het bekijken van de website. Dit principe zorgt ervoor dat de website voor de gebruiker meer aanvoelt als een applicatie dan een website.

¹ Een website op een server uploaden, zodat hij gekoppeld kan worden aan het internet. [4]

Het probleem

Als je een website online zet is hij bereikbaar via zijn URL, maar je zal hem nog niet vinden als je gaat zoeken via een zoekmachine. Hiervoor moet hij eerst vindbaar gemaakt worden voor zogenoemde crawlers¹ (spiders). Zodra de crawlers een pagina vinden, maken ze vaak een lokale kopie om deze later te verwerken en te indexeren voor bijvoorbeeld zoekmachines. Een MPA wordt op deze manier volledig geïndexeerd en alle content kan worden gevonden, omdat elke pagina "statische" content heeft. Bij een SPA is dat een ander geval. Je haalt namelijk geen 'kant en klaar' nieuw document op van de server, maar je verandert alle content dynamisch binnen het huidige document. Een crawler weet op deze manier niet wanneer er aanpassingen gedaan worden. Hierdoor word je website niet goed gecrawled en geïndexeerd. En dat is een groot probleem voor de SEO-score van je website. Doordat je score dan zo laag is zal je ook zeker niet hoog in de zoekmachines verschijnen.

4

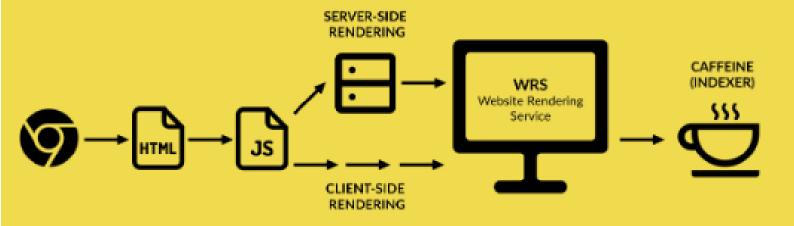
¹ Een bot die het wereldwijde web op een methodische en geautomatiseerde manier doorbladert.

Onderzoeksvraag

Hoe kunnen we de SEO van Single Page Applications verbeteren en wat zijn daarin de belangrijke factoren?

Deelvragen

- Hoe wordt een website gerenderd?
- Wat is Google crawling, indexing en SEO-ranking?
- Waar wordt de SEO-score op gebaseerd?
- Welke SEO-technieken kunnen we anders of beter toepassen voor SPA's?



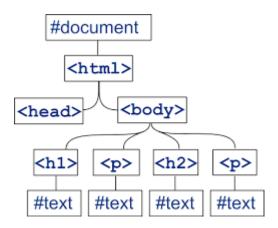
Rendering-banner. (z.d.). [Foto]. Geraadpleegd van https://i0.wp.com/seotradenews.com/wp-content/uploads/a-survival-kit-for-seo-friendly-javascript-websites.png?resize=559%2C250&ssl=1

Hoe wordt een website gerenderd?

Om uit te leggen hoe een website gerenderd wordt moeten we eerst duidelijk hebben wat rendering-inhoud. Het omzetten van html tags binnen de pagina naar bepaalde commando's voor het OS, waardoor alle content van een pagina in goede orde geladen wordt, heet renderen.

Wat doet daarin de rendering engine?

Om te beginnen moeten we eerst weten wat een rendering engine doet. Een rendering engine zorgt ervoor dat er iets op je scherm verschijnt op het moment dat je een website bezoekt. Hoe dit in zijn werk gaat is uit te leggen aan de hand van een stappenplan.



1. HTML parsen en de DOM tree creëren

Er wordt een structuur opgezet uit vanuit HTML. Zo bestaat een pagina normaal gesproken uit een beginnende httml> tag met daarin vaak head> en body> tags. Daarin worden dan alle andere elementen aangemaakt. Dit proces kan zo door blijven gaan en dat wordt nesten genoemd. Als alle elementen aangemaakt zijn door de rendering engine krijg je een soort boom, deze boom wordt de "DOM tree" genoemd. Het is dus uiteindelijk een soort boom gemaakt van HTML, waar elke tag een tak representeert die begint vanuit het root element (de https://doi.org/

2. Het bouwen van de Render tree

Nadat de DOM tree is aangemaakt, wordt de CSS geparst en toegevoegd aan de DOM tree. Alles samen wordt de Render tree genoemd. Deze Render tree bevat nu visuele elementen zoals de hoogte, breedte en kleur van de elementen, maar wel in dezelfde volgorde als hoe ze in de browser getoond worden.

3. Het layout proces

Als de Render tree gevormd is zal de render engine systematisch door de HTML-elementen lopen om uit te zoeken waar ze op het scherm geplaatst moeten worden. Standaard begint dat op 0, 0 in de linkerbovenhoek. Alle elementen en attributen worden op specifieke coördinaten geplaatst.

4. Inkleuren

De laatste stap is eigenlijk het volledig "tekenen" (renderen) van elke tak binnen de Render tree. Dit wordt gedaan door middel van communicatie met het OS-Interface waar alle designs en stijlen in zitten voor de UI-elementen. Zo wordt alles gestructureerd op je beeldscherm getoond.

Om het hele render-proces te starten moet er eerst een aanvraag gedaan worden naar de server. Dat gebeurt normaal gesproken door middel van een click event, denk dan bijvoorbeeld aan het klikken op een link. Zodra je klikt gaat er een aanvraag naar de server om de benodigde data op te halen voor de nieuwe pagina. Hierbij maken we gebruik van het "Hypertext Transfer Protocol" 1. Zodra er een response terugkomt vanaf de server gaat de rendering engine aan het werk. Elke keer als je dus op een link, of iets soort gelijks, klikt waarna je naar een nieuwe pagina gaat wordt dit hele proces herhaald. Hierdoor kan het zijn dat je als gebruiker "even" moet wachten tot de pagina volledig is geladen.

[7]

Hoe wordt een SPA gerenderd?

Een SPA kan op verschillende manieren worden gerenderd. Elke manier heeft zo zijn voor en nadelen. Je hebt Client-side rendering en Server-side rendering. In essentie gaat het erom dat je één pagina hebt die verschillende content/data kan bevatten. De DOM-elementen op de pagina zijn afhankelijk van deze data.

Client-side rendering (CSR)

Met CSR begin je met een 'lege' pagina in de browser. Dan wordt er een call gedaan naar de server voor de benodigde data op die pagina. Vervolgens wordt dan alle data vanuit de server teruggestuurd en kunnen de DOM-elementen weergegeven worden. Als laatste worden dan de CSS en JavaScript er nog aan toegevoegd en is de pagina klaar voor gebruik.

Met deze manier van renderen hoeft de server vrij weinig te doen. Hij hoeft namelijk alleen de data te sturen wanneer nodig. Het nadeel hiervan is dat je alles in de browser van de client doet en daar kunnen crawlers niks mee. Hierdoor word je website niet of nauwelijks geïndexeerd en hoef je al helemaal niet meer te denken aan SEO. Een wat betere optie zou Server-side rendering zijn.

Server-side rendering (SSR)

Bij SSR gaat het net iets anders. Hier wordt namelijk op de server al de pagina met alle DOMelementen opgebouwd aan de hand van de verkregen data. Nu heb je een 'ready to be rendered' HTML-pagina en deze wordt dan teruggestuurd naar de browser. Daar wordt dan de CSS en JavaScript er nog aan toegevoegd en is de pagina klaar voor gebruik.

Dit proces wordt normaal binnen een paar milliseconden uitgevoerd, maar dat is niet altijd het geval. Het ligt bijvoorbeeld aan de snelheid van je internet, de locatie van de server of hoeveel mensen tegelijk dezelfde website proberen te bezoeken hoe snel alles geladen wordt.

Prerendering

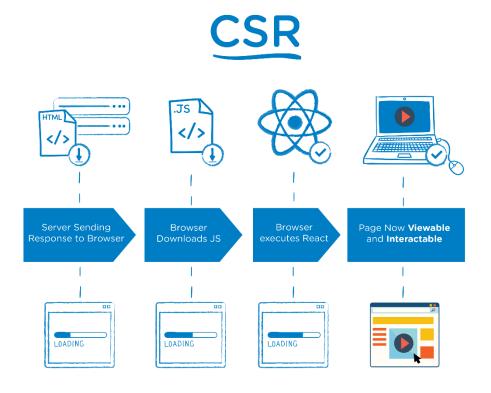
Een extra optie is nog prerendering. Wat je hiermee doet is alle elementen op een pagina van tevoren laden en cachen, zodat een crawler ze direct kan zien. Een prerender service is in staat een pagina request te onderscheppen en vervolgens te checken of het vanuit een crawler of vanuit een normale gebruiker komt. Als het request vanuit een crawler komt wordt namelijk de gecachte versie van de website getoond en anders wordt alles normaal geladen.

Bronnen: [12]

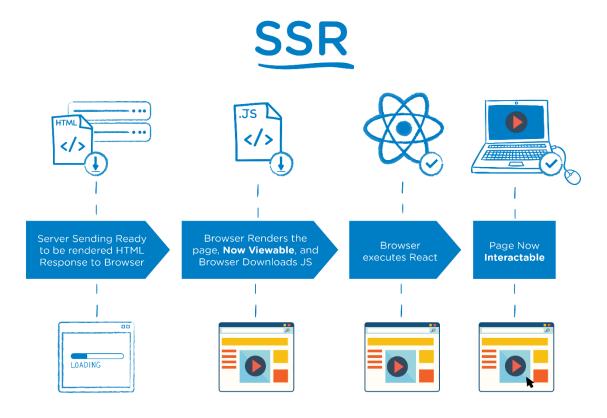
8

¹ De manier waarop browsers en servers met elkaar communiceren.

Even vergelijken



Overzicht Client-side rendering. (2017.). [Foto]. Geraadpleegd van https://miro.medium.com/max/4465/1*CRiH0hUGoS3aoZalY4H2yg.png



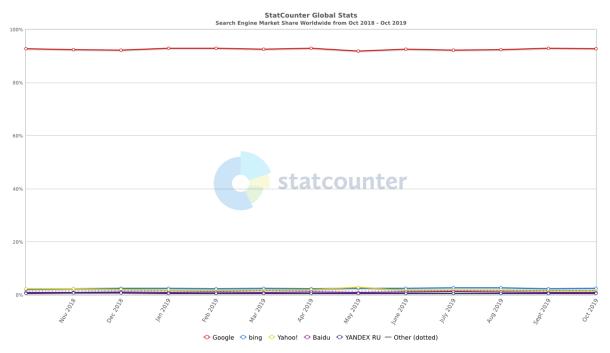
Overzicht Server-side rendering. (2017). [Foto]. Geraadpleegd van https://miro.medium.com/max/4234/1*jJkEQpgZ8waQ5P-W5lhxuQ.png



SEO-banner. (z.d.). [Foto]. Geraadpleegd van http://www.allininfosystems.com/wp-content/uploads/2016/03/seo_new_banner.png

Wat is Google crawling, indexing en SEO-ranking?

Dus nu weten we wat een website is, hoe hij online komt te staan en hoe ze gerenderd worden. Het volgende probleem is nu, hoe kunnen mensen hem vinden en bezoeken zonder de directe URL in te hoeven voeren? Om deze vraag te beantwoorden hebben we een aantal dingen nodig, namelijk: crawling, indexing en daarbij ook SEO-ranking. Dit wordt allemaal gedaan voor de zoekmachines. Om dit deel schaalbaar te houden heb ik ervoor gekozen Google als zoekmachine te nemen. Google is de grootste en meest gebruikte zoekmachine wereldwijd.



StatCounter-Global-Stats. (2019). [Foto]. Geraadpleegd van https://gs.statcounter.com/search-engine-market-share

Google crawling

Om je website te laten crawlen door Google moet je als website-eigenaar een aantal dingen verstrekken. Je moet sitemaps en een robot.txt aanleveren. Zo weten crawlers zoals Googlebot wat er wel en niet gecrawled moet worden op jouw website. Vervolgens gaan de crawlers aan de slag. Wat ze doen is jouw website bezoeken en alle links op jouw website afgaan om nieuwe pagina's te zoeken. Een crawler ziet dan jouw website alleen heel anders. De website van de Hogeschool Rotterdam ziet dan ongeveer zo uit:





Er wordt door een crawler specifiek gekeken naar een aantal punten zoals: Links, H1's, H2's, Meta-teksten (etc.). Die punten spelen een grote rol als het gaat om SEO. Ook gaat de software opzoek naar wijzigingen in bestaande links en zoekt het naar kapotte links. Alle info hierover sturen ze dan terug naar de Googleservers waar de data verder verwerkt kan worden.

Google biedt daarbij zelf ook uitgebreide programma's aan waar je als website-eigenaar inzicht en controle hebt over een deel van dit proces. Zo kan je met Search Console zelf heel gedetailleerd aangeven hoe Google jouw site moet crawlen. Je kan je aan- of afmelden voor crawls, je kan aangeven welke delen ze wel of niet mogen crawlen en je kan de resultaten van crawls bekijken. Kortom ze bieden je genoeg mogelijkheden om te zorgen dat Google jouw website goed interpreteert.

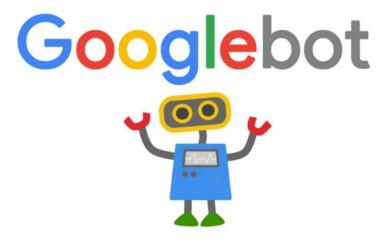
Google indexing

Het indexeren gebeurt allemaal aan de kant van Google. Als een crawler dus een webpagina vindt, wordt alle content ervan weergegeven op de minimale manier als eerder besproken. Ze checken daarna de pagina op belangrijke punten zoals H1's, zoekwoorden, hoe recent de website eigenlijk is en meer. Al deze dingen worden uiteindelijk bijgehouden in de zoekindex van Google. De Google Zoeken-index blijft continu doorgroeien hierdoor, maar blijft dus ook up-to-date.

Google SEO-ranking

Google's SEO-ranking is eigenlijk heel simpel gezegd een ranglijst waar jouw site in komt te staan op basis van de bijbehorende SEO-score. Hoe beter je score, hoe hoger je in de ranglijst komt te staan. Je wil natuurlijk altijd de hoogst mogelijke score halen, want dan kom je op de eerste pagina te staan op Google.

Waarom is die eerste pagina zo belangrijk dan? Nou, heel simpel, als je iets gaat zoeken op Google wil je het zo snel mogelijk vinden. Je wilt niet blijven zoeken naar dat ene ding wat je nodig hebt. Alles wat dus bij Google op de eerste pagina staat heeft meer kans om bekeken te worden.



Googlebot. (z.d.). [Foto]. Geraadpleegd van https://www.miltonmarketing.com/news/googlebot-runs-latest-chrome-this-is-big-news/
Bronnen: [9], [10]



SEO-score. (z.d.). [Foto]. Geraadpleegd van http://chescoweb.com/wp-content/uploads/2018/07/SEO.png

Waar wordt de SEO-score op gebaseerd?

De SEO-score wordt opgebouwd uit scores vanuit een aardig aantal punten. Deze scores wegen niet allemaal even zwaar mee, maar ze zijn uiteindelijk wel belangrijk voor de totaalscore.

Technical (Hoge prioriteit)

Het technische deel gaat vooral over de bereikbaarheid van elementen. Denk dan vooral aan hoe de code structuur van de website is, het toevoegen van een robots.txt voor de crawlers, een sitemap en tegenwoordig wil je ook rekening houden dat je website gebruik maakt van HTTPS. Want Google houdt er wel een klein beetje rekening mee volgens een tweet van John Mueller (Webmaster Trends Analyst at Google):

Yeah, that's wrong. HTTPS is not a factor in deciding whether or not to index a page, at all. We do use HTTPS as a light-weight ranking factor, and having HTTPS is great for users. A free certificate from Let's Encrypt works just as well. (Mueller, 29 januari 2019, Twitter)

Als dit namelijk niet mogelijk is zal je website nooit goed geïndexeerd worden en zal vervolgens ook niet in de zoekmachines verschijnen. Daardoor komen er ook geen bezoekers en zal je dus conversie mislopen.

Zodra er dan een Technische SEO Audit gedaan wordt komen al deze punten vanzelf boven. Daarom wordt vaak in de Development fase al nagedacht over de meeste SEO-punten, zodat die al vanaf begin af aan aanwezig zijn binnen de website. Het is nou eenmaal niet zo handig om alles achteraf te gaan regelen wat SEO betreft. Ook kan je voor sommige CMS-systemen gebruik maken van Plugins die de SEO voor je regelen (bijv. Yoast voor WordPress).

Content (Medium prioriteit)

Onder content vallen de meta-content sets, opbouw van teksten, semantische HTML-elementen (<section>, <header>, <article>, <blockquote> etc.), keywords en nog veel meer. Veel van de content is alleen op de pagina's zelf te vinden, maar een aantal dingen zie je al terug in de zoekmachine resultaten. Als we bijvoorbeeld even kijken naar de titel en Meta-teksten, dat zijn de onderdelen die je direct ziet als je een zoekopdracht hebt gedaan in een zoekmachine. Dat ziet er dan uit als volgt:

Hogeschool Rotterdam - Overtref jezelf met een van onze ...

https://www.hogeschoolrotterdam.nl -

Hogeschool Rotterdam biedt ruim 80 bacheloropleidingen en Associate-degreeprogramma's aan. Het aanbod bestaat verder uit contractonderwijs, post-hbo en ...

Meta tekst

Het is dus ook belangrijk dat de titel en Meta-teksten aansluiten op de content van je website. Daarmee voorkom je dat de gebruiker je website bezoekt, ziet dat je website content totaal niet overeenkomt met je meta teksten, en dus direct teruggaat naar de zoekmachine.

User Experience (Lage prioriteit)

Dit gaat om hoe effectief embedded informatie zoals: links, animaties en beeldmateriaal gebruikt worden om de user experience te verbeteren. Voor de SEO-score maakt het niet uit hoe mooi je animaties zijn of hoe scherp je foto's zijn. Je hebt een gestructureerde link hiërarchie nodig met foto beschrijvingen om hem indexeerbaar te maken en daarnaast ook de usability te verbeteren.

En om de usability op je website te waarborgen zijn er richtlijnen opgezet. Zo heb je A11Y, ook wel bekend als het Accessibility Project, wat zich hier veel mee bezig houdt. Zij hebben op checklists opgesteld die je af kan gaan om te kijken of jouw website voldoet aan de A11Y richtlijnen. Verder heb je ook nog WCAG, Web Content Accessibility Guidelines, dit is een document wat is opgesteld door het World Wide Web Consortium¹ (W3C). Dit document bestaat voor een groot deel uit richtlijnen voor de toegankelijkheid van content op he web, met oog op mensen met een beperking.

Mobile (Medium prioriteit)

Als laatste is er ook nog mobile. Mobile wordt steeds meer het primaire apparaat waar websites op bezocht/bekeken worden. Hiervoor moet de website wel in staat zijn op mobile getoond te worden. Om hier goed op te scoren worden websites responsive gemaakt, dat houdt in dat ze mee schalen naar de grootte van het scherm. Maar dat is niet het enige wat er moet gebeuren, want je moet nadenken over andere functionaliteiten in verband met het touchscreen.

Bronnen: [11]

14

¹ Organisatie achter de ontwerpen van webstandaarden op het world wide web

Welke SEO-technieken kunnen we anders of beter toepassen voor SPA's?

Om te beginnen kan is het goed om van tevoren al rekening te houden met SEO in je SPA. Het makkelijkste is om dit te doen door alle content standaarden (denk aan H1's, links, meta-tekst alt-tekst etc.) mee te nemen in het development proces. Op die manier kan je van begin af aan een goede structuur opzetten. Wat daarbij ook handig is om de A11Y en WCAG-richtlijnen daarbij te betrekken, zodat je alles ook opbouwt volgens de meest gebruikte web standaarden qua content. Zodra je alles voor de content van je SPA geregeld hebt is het tijd om te gaan kijken naar CSR, SSR en prerendering.

Het beste is om ten minste te kiezen voor CSR in combinatie met SSR. Door de SSR kunnen de crawlers je website crawlen en indexeren. De CSR zorgt dan nog steeds voor de dynamiek van de content en draagt daarmee bij aan een soepele user-experience.

Als je ervoor kiest om een SPA te bouwen met React JS kan je gebruik maken van Next JS. Next JS is een React Framework waarmee je SEO vriendelijke SPA's kan bouwen. NEXT JS heeft SSR geïntegreerd, hierdoor hoef je zelf niet te veel tijd te spenderen aan het opzetten van SSR, maar kan je het makkelijk toevoegen aan je SPA. Er bestaat ook eenzelfde Framework voor Vue en dat heet NUXT JS. Het heeft precies dezelfde functionaliteit als NEXT JS, maar het is bedoeld voor op Vue gebaseerde applicaties.

Je kan voor je React SPA ook gebruik maken van de Redis Plugin. Hiermee kan je je website cachen op de server. Deze gecachte website kan dan vervolgend door crawlers ook gebruikt worden en kan je website toch gecrawled en geïndexeerd worden.

Conclusie

In dit onderzoek is er gekeken hoe we de SEO van Single Page Applications kunnen verbeteren en wat daarin de factoren zijn. Er is gebleken dat je een aantal dingen nodig hebt om de SEO voor je SPA te verbeteren. Zo heb je in ieder geval een manier nodig, zodat crawlers je volledige SPA kunnen crawlen. Dit kan je oplossen met onder andere SSR en prerendering. Om het jezelf iets makkelijker te maken kan je daarvoor gebruik maken van Frameworks en/of Plugins (bij. Next JS, NUXT JS, Redis etc.) Verder heb je ook nog bepaalde richtlijnen voor webcontent die je moet volgen waardoor je content voor een betere SEO-score zal zorgen. En als laatste heb je de standaard punten waar bijvoorbeeld Google op checkt, waar vervolgens je SEO-score op wordt gebaseerd. Al deze punten samen zouden ervoor moeten zorgen dat je SPA SEO verbeterd wordt.

Bronnenlijst

Wetenschappelijke bronnen

- [1] Tamjidi, S. (ter perse). Competition to get the best ranking on SEO. Comparison between SPA and MPA. Geraadpleegd van http://www.diva-portal.org/smash/get/diva2:1143657/FULLTEXT02
- [2] Ghose, A., Ipeirotis, P. G., & Li, B. (2014). Examining the Impact of Ranking on Consumer Behavior and Search Engine Revenue. Management Science, 60(7), 1632–1654. https://doi.org/10.1287/mnsc.2013.1828
- [3] Malaga, R. A. (2008). Worst practices in search engine optimization. Communications of the ACM, 51(12), 147. https://doi.org/10.1145/1409360.1409388

Overige bronnen

- [4] Wat is hosting? | SoHosted. (2018, 13 juli). Geraadpleegd op 22 november 2019, van https://www.sohosted.com/webhosting/wat-is-hosting/
- [5] Website. (2019, 11 november). Geraadpleegd op 22 november 2019, van https://nl.wikipedia.org/wiki/Website
- [6] Blacquière, E. (2019, 24 september). Wat is SEO? Uitleg & Werking EdWords.nl Search Marketing Blog. Geraadpleegd op 22 november 2019, van https://www.edwords.nl/kennisbank/definitie/seo/
- [7] Website. (2018, 21 november). What Does It Mean To "Render" a Webpage? Geraadpleegd op 22 november 2019, van https://www.pathinteractive.com/blog/design-development/rendering-a-webpage-with-google-webmaster-tools/
- [8] Single-Page Websites: Are They Good or Bad for SEO? (z.d.). [Foto]. Geraadpleegd van https://www.link-assistant.com/news/one-page-website-seo.html
- [9] Hoe Google Zoeken werkt | Crawlen en indexeren. (z.d.). Geraadpleegd op 22 november 2019, van https://www.google.com/intl/nl/search/howsearchworks/crawling-indexing/
- [10] Googlebot Search Console Help. (z.d.). Geraadpleegd op 22 november 2019, van https://support.google.com/webmasters/answer/182072?hl=en
- [11] Needham, S. (2019, 21 november). *What is the SEO Score?* Geraadpleegd op 22 november 2019, van https://support.siteimprove.com/hc/en-gb/articles/115001754792-What-is-the-SEO-Score-
- [12] Breux, G. (2018). Client-side vs. Server-side vs. Pre-rendering for Web Apps. Geraadpleegd van https://www.toptal.com/front-end/client-side-vs-server-side-pre-rendering