SEO AUDIT REPORT

[REDACTED]

April 2017

Prepared by Deep Field, Inc.

Overview

Sources

The goal of this audit is to evaluate www.[URL REDACTED] and two subdomains (blog.[URL REDACTED] and press.[URL REDACTED]) for health in a variety of areas in order to determine how the site might be better optimized for search engines and end users. While the focus of the audit is [URL REDACTED] and the two subdomains, we crawled all subdomains that were open in order to understand the overall structure and how it might impact the SEO performance of the root domain and two subdomains. In order to complete the audit the following sources were reviewed:

- Google Analytics profile
- Google Search Console
- Independent site crawls
- Data collected via 3rd party sites such as MOZ, SEMRush, Ahrefs, etc.

Report Organization

This report describes the findings of the audit and related recommendations, action items and follow-up questions that resulted.

- **Recommendations** - These are suggestions and proposals that Deep Field, Inc. would undertake if they were managing the campaign for the site. Typically, these are not specific action items, but instead are general guidelines for optimization and performance.
- **Action Item** - These are specific actions/tasks that need to be taken immediately.
- **Questions** - These are questions that emerged during the audit process and would be valuable for you to answer

General Information

The domain was initially registered in 2007 and recognizable captures by Wayback.archive.org can be found going back to 2014. The domain registrar is Gandi SAS (www.gandi.net) with hosting by Amazon and nameserver provided by Amazon Route 53. The domain is registered to [REDACTED].

The site ([URL REDACTED]) has an Alexa Traffic Rank of 24,043 (Global) and 3,359 (United States). Alexa engagement metrics show a bounce rate of 35.5%, daily pageviews per visitor as 2.77 and daily time on site as 3:15 minutes.

Google Analytics data (All Website Data View) is limited, showing only for the month of April 2017. From April 1 to April 25 Google Analytics for www.[URL REDACTED] recorded:

- 388,966 sessions
- 73,722 users

www.[URL REDACTED]
from Wayback Archive 2014

- 2,800,724 pageviews
- 7.20 pages per session
- 00:25:28 average session duration
- 8.92% bounce rate

Of these sessions, Language EN-US accounted for 99.23% and all other languages were negligible. Nearly all sessions (98.47%) were from the United States along with Canada (.29%), United Kingdom (.17%), Netherlands (.14%), Australia (.11%), India (.10%), Taiwan (.05%), Singapore (.04%) and Spain (.04%).

Google Analytics traffic acquisition data looks mostly normal, though it is uncertain if there is some overlap between the referral and direct traffic based on it coming from [REDACTED] subdomains. Interestingly, most of the referral traffic is from [REDACTED] subdomains except for a small portion coming from google.com, Google sites and Symbaloo. The Google Sites and Symbaloo are natural referral sources since [REDACTED] are likely using them as bookmark pages. Two other notable issues are the high bounce rate (42.78%) on paid search acquisition and the very low amount of social traffic.

- Referral (139,047)
- Direct (134,374)
- Organic Search (87,916)
- Other 24,056
- Paid Search (1772)
- Social (582)
- Email (25)

Google Analytics data (blog.[URL REDACTED]) for the [REDACTED] blog subdomain shows no visits until April 14th. For the dates between April 14 and April 25 Google Analytics recorded:

- 5,598 sessions
- 4,535 users
- 13,793 pageviews
- 2.46 pages per session
- 00:02:54 average session duration
- 4.0% bounce rate

Of these blog sessions, EN-US Language accounted for 86.25% and EN-GB for 5.75%, with other languages being negligible. Most of the traffic was based in the United States (4,019 sessions / 71.79%).

Google Analytics traffic acquisition data for the blog shows a higher percentage of organic search, though still very little from social:

- Organic Search (4,002)
- Direct (941)
- Referral (341)
- Social (254)
- Other (50)
- Email (7)
- Paid Search (3)

Referral traffic for the blog has a high bounce rate which appears to be a result of an 81.3% bounce rate from app.hubspot.com.

SEO Audit

Technical SEO

This section of the report focuses on technical aspects, including accessibility and indexability of the site in order to determine if there are any issues negatively impacting user experience or search engine performance.

Observations
- The site's structure is complex and there are a significant number of subdomains used.
- A significant percentage of the site's pages are deep, requiring many clicks in order to access (42.55% have a click depth of 8 or more).
- There are redirect chains on the root domain and several subdomains. Moz currently located 18 of these, but independent crawls suggests there are more. Each of the redirects involve three pages (2 hops) and each hop typically reduces authority and increases load times for the end user.
- The current robots.txt file (https://www.[URL REDACTED]/robots.txt) is functioning correctly and is set to 'full allow' (all content can be crawled). The previous robots.txt was returning a 403 response, informing all crawls that the site was forbidden. Any non-public elements of the site should be disallowed in the robots.txt file. The sitemap.xml location is not currently included in the robots.txt file.
- The sitemap.xml files (https://www.[URL REDACTED]/sitemap.xml / http://blog.[URL REDACTED] / http://press.[URL REDACTED]) are setup and functioning correctly. If there are plans to significantly scale the site in the future (i.e. multiple languages/international) then it might be helpful to consider a sitemap index file, but it is not necessary at this time.
- There is limited data in Google Webmaster Tools as the property is for the naked domain https://[URL REDACTED], but there is nearly no indexation (only 6 pages) and no sitemaps have been submitted.
- The SSL certificate is installed and working correctly, though it is set to expire in July of this year. Most resources on the root directory have been redirected to https, though some subdomains are not secured.
- There is a miscellaneous 403 Forbidden Access returned on https://www.[URL REDACTED]/undefined
- Moz found 28 broken internal links (returning 404 not found error codes). Most of these were on subdomains, though some are for public content.
- There were no canonical tags pointing to redirects.
- The only URLs no indexed are on the ideas.[URL REDACTED] subdomain
- There is very limited use of schema/markup data on www.[URL REDACTED], press.[URL REDACTED] and blog.[URL REDACTED]. None shows in Google Webmaster Tools (likely a result of the property setup issue) and tests in Google's Structured Data Testing Tool only returned evidence of markup for video objects on www.[URL REDACTED].
- Google PageSpeed Insight tool returned a score of 79/100 for www.[URL REDACTED] on desktop. Mobile, however, returned a lower score of 57/100.
- Google PageSpeed Insight tool returned a score of 40/100 for blog.[URL REDACTED] on desktop and 49/100 on mobile.
- Google PageSpeed Insight tool returned a score of 79/100 for press.[URL REDACTED] on desktop and 57/100 on mobile.

- The URL structure throughout the site and subdomains generally looks good. There are some inconsistencies with the use of dashes and underscores, capital letters, and parameters, but most of these are related to the ideas.[URL REDACTED] subdomain so have no impact on SEO. Additionally, there are 587 URLs with a long length over 115 characters, but they also are on the ideas.[URL REDACTED] subdomain.

Recommendations
- Move everything, including subdomains, to https.
- Investigate how and why redirect chains are being created so that processes can be established to curb their creation in the future.
- Based on the Google PageSpeed Insights, consider additional optimization of images on www.[URL REDACTED] and whether browser caching could help with the mobile speed score/experience. The same considerations should apply to press.[URL REDACTED]. Both the mobile and desktop results were slower for blog.[URL REDACTED] so it is recommended to further investigate the cause.

Action Items
- Setup Google Webmaster Tools and ensure that all sitemap.xml files have been submitted and that indexation is showing up in the interface.
- Reduce the number of subdomains on the site by moving content to subdirectories on the root domain. In most cases, when you are developing public content that benefits SEO it should reside on the root domain instead of a subdomain. Ideally, blog.goguaridan.com would be moved to www.[URL REDACTED]/blog. At the same time, we understand that you are using HubSpot and don't necessarily have control over the use of a subdomain.
- Resolve the redirect chains on the site
- Include disallows in the robots.txt for any non-public part of the entire [URL REDACTED] domain. This would include things like licenses, management paths, user paths, theme assets, admin/login/logout/password reset pages, scripts, etc. The robots.txt that was on https://[URL REDACTED] had disallows for a few of these, but it should be extended in order to help with crawl budget. Additionally, whether choosing to use https://[URL REDACTED] or https://www.[URL REDACTED] for the root domain, include identical robots.txt files on both.
- Include the sitemap.xml address in the robots.txt file. It's best to submit the sitemap.xml to Search Engines, but including it in the robots.txt will ensure it is recognized by all, even for those that you do not manually submit.
- Resolve the 404 errors on the site
- Implement the use of schema/markup data, focusing especially on blogs and press articles.
- Address the high bounce rate from referral traffic on blog.[URL REDACTED].

Questions
- What factors and strategies influence the creation of many subdomains in the site structure/architecture?

On-Page SEO

This section of the report focuses on characteristics of the site's pages that impact its rankings in search engines, including content and HTML source code that can be optimized.

Observations

> These observations are specific to the root domain, press and blog unless otherwise stated. There are on-page issues with other subdomains, including some that are public content that impact SEO, but are excluded from the scope of the comments below.

- Generally, it appears that there is little to no keyword strategy used throughout the root, press subdomain and blog subdomain except for select keywords on the front sales pages (home/admin/teacher/fleet).
- There are some pages (Moz identified 7 so far and independent crawls identified 8) with duplicate page content. So far the duplicate pages that have been identified are ones that will likely be resolved when disallows are added to the robots.txt.
- There are pages that are identified as having long page titles (over 56 characters or over 568 pixels). This impacts the way that these display in SERP results when users search.
- Press.[URL REDACTED] is missing a meta description
- There are a significant number of pages (131) with duplicate meta descriptions.
- Meta description for blog authors are page counts are being read as either too long or too short.
- There are four images missing alt text.

Recommendations
- Setup an ongoing review cycle to evaluate old content on the site that is stale and underperforming. Remove the old content that is receiving little traffic and consider revising it for better performance.

Action Items
- Develop keyword strategies for each content segment throughout the site to maximize reach and ensure minimal cannibalization.
- Address the long page title issues
- Add a meta description for press.[URL REDACTED]
- Address the duplicate meta description issues
- Address the meta description length issues
- Address the missing alt text in four images

Questions
- What processes are you using to identify content development needs? How are you testing/evaluating content to see how it is performing?
- How are on-page SEO elements being tested and optimized? How do you know when they are working?

Off-Page SEO

This section of the report focuses on off-page factors that impact how search engines determine the site's trustworthiness and authoritativeness. Examples of factors that impact a sites trustworthiness and authoritativeness are backlink quality and social media sharing.

Observations
- The MOZ Domain Authority (DA) for [URL REDACTED] is 44/100 and a MozTrust rating of 5.45/10. Moz also rates the Domain Authority for hapara.com as 44/100, though lightspeedsystems.com is slightly higher at 46/100 and iboss.com tops the list with 50/100. Domain Authority is a score that helps predict the overall health of your site in relation to search engine ranking. The current score of 44/100 is a decent score for a site of the age and size of [REDACTED], especially considering the SEO work to date.
- Ahrefs shows that 30,000+ links from nearly 700 referring domains. There has been relatively steady growth of referring domains since 2014 outside of a couple of dropoffs in February/March 2015 and June 2016, though these were both relatively small losses. Interestingly, there is an unusually low number of .gov and .edu referring domains considering the nature of the [REDACTED] product and the clients it serves. There are three .edu referring domains (havant.ac.uk, universityofcalifornia.edu, univ-paris5.fr) and no .gov referring domains.
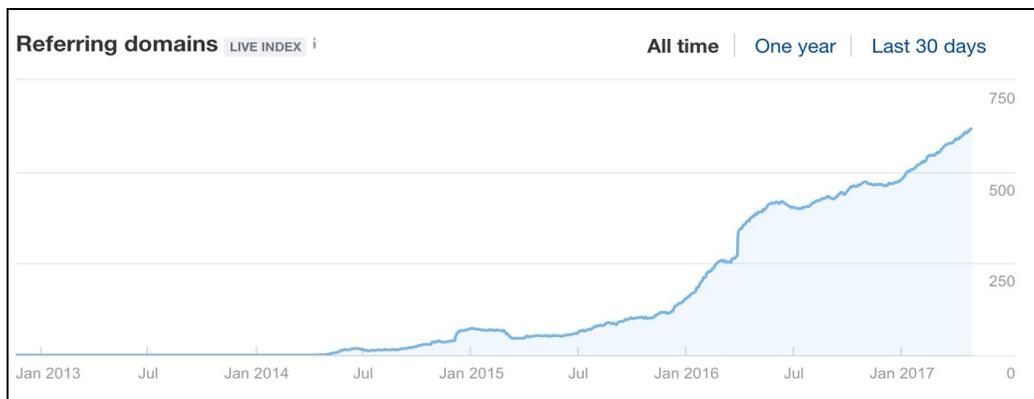
- There are four backlinks that are linked to pages returning 404 not found errors.
- There are 10 outgoing links which refer to pages returning 404 not found errors.
- There were no spammy links found in the Moz inbound link report
- The social media engagement with the [REDACTED] brand is minimal. This is especially noticeable in when comparing direct competitors and other online educational tools with higher Domain Authority.

Enter your domains	1	2	3	4	5	Compare
Domain Rating	51	51	57	61	**65**	
Facebook	2	336	416	9,539	**88,102**	
Google+	1,889	45	17	2,643	**78,517**	
LinkedIn	29	256	16	10	**742**	
Pinterest	0	0	0	0	**36**	
Ref Domains	617	546	1,102	6,865	**21,711**	
Ref .gov domains	0	0	2	13	**203**	
Ref .edu domains	2	10	9	210	**624**	
Ref .com domains	276	307	658	3,923	**14,292**	
Ref .net domains	34	19	90	351	**839**	
Ref .org domains	188	83	120	1,294	**2,286**	
Ref IPs	494	482	972	4,085	**9,552**	
SubNets	425	398	833	3,193	**7,004**	
Referring Pages	31,404	12,502	25,960	1,990,060	**3,055,647**	
Backlinks	32,271	25,526	29,260	**4,925,398**	4,429,541	
Crawled Pages	2,142	2,170	17,988	37,723	**4,964,271**	
Text	30,967	25,366	28,992	**4,866,317**	4,371,992	
DoFollow	30,283	19,821	27,009	**4,841,385**	4,163,498	
NoFollow	684	5,545	1,983	24,932	**208,494**	
Redirect	1,243	45	201	2,094	**3,739**	
Image	362	6,479	407	159,936	**1,360,714**	
Form	0	0	1	**3,807**	9	
.gov	0	0	5	64	**137,594**	
.edu	9	45	44	**169,597**	118,205	

Recommendations

- A comprehensive keyword research project along with a content gap analysis (between [REDACTED], direct competitors and growth competitors) is recommended in order to guide development of new content.
- Develop strategies for targeting .gov and .edu backlinks as these type of links have an especially authoritative value and [REDACTED] is well-positioned to obtain these. They need to be connected to the root domain. One possible strategy would be to develop a teacher and/or director ambassador program and expecting a backlink as part of the agreement with each ambassador.

Action Items

- Address the four backlinks and 10 outgoing links that reach pages returning 404 not found errors.
- Engage students, parents, teachers, school administrators, district leaders and governmental leaders in social media.

Questions

- None

Addendum

Various reports used or produced to conduct the audit are included.