100 Queries: What do Dutch Children See on the Web?

Carsten Schnober[1,*], Maarten Sprenger[2]

[1]*Netherlands eScience Center, Science Park 402 (Matrix THREE), 1098 XH Amsterdam, The Netherlands*
[2]*Slim Zoeken, BK49 Tweede Boerhaavestraat 49, 1091 AL Amsterdam, The Netherlands*

Abstract

Using a random sample of authentic queries, we have employed Google Search as a lens to get a representative view on what Dutch children see on the web. With manual expert knowledge, we have annotated in-depth information about individual result pages as well as about the underlying sources.

Our data reveals strong correlations between the intents of sites that are relevant to Dutch children, and factors including the 'About' page quality and the number of ads. To facilitate further analysis, we make the raw data available along with our annotations.

We also propose a generic methodology using a query sample in combination with a search engine to draw conclusions about content that is available and visible on the web for a specific audience. This method is applied to empirically approach questions such as: Does the web contain relevant content and sufficient coverage for a specific target group? Which are prominent sources for available information, and what is their quality? Who is behind a source, and is such information transparent?

Questions for information retrieval practitioners follow: Is a generic web search engine beneficial for an audience, or should they directly navigate to the known credible sources for their needs? Should they use domain-specific or site-internal search engines, while broader web search engines are rather a distraction on the way to towards gaining information in the most effective way?

Keywords
information retrieval, evaluation, education, web search, annotation

1. Introduction

The dynamics of the web make it hard to estimate, even for experts: what do internet users get to see "in the wild"? Existing pages change constantly, new sources arrive, while search engines update their indices and their algorithms. From the practical point of view in this work, we are therefore less interested in what content exists anywhere on the web, but rather in which part of the web is visible to a particular audience. And consequently: does the web contain the content that is relevant for a specific audience, and is it accessible with current tools?

For the purpose of educating children and teachers in the Netherlands on digital literacy and information access tools in particular, we have asked ourselves the following questions: Which sites do Dutch children encounter on the web? How many of them are trustworthy and useful? What are their intents in terms of information provision and commercial interests?

We have acquired and refined a sample of authentic queries used by Dutch children between 8 and 12 years old. Using Google search results for these queries yields a sample of web sites that are visible for the audience. We have analysed these results and the underlying sources manually and developed a taxonomy covering categories like intent, target age group, and accessibility. The combination of authentic queries and search engine results serves as a vehicle to understand which content actual users get to see.

Analysing the sources behind the results provides a perspective on the motivation and general credibility of specific sites, but also about the information landscape as a broader context for individual pages and sites. Our in-depth analysis of sources allows us to draw conclusions about what types of sources the members of this group are expected to encounter in general.

All our data and code used for analyses are publicly available[1] and we encourage the community to use for further analyses and/or extend it.

2. Related Work

Trustworthiness and credibility of a web page involves various factors, and has been a topic of research for decades [1, 2]. Aspects such as information about the author(s) of a page, page structure and design, linguistic properties play a role, in particular when it comes to disputed topics [3] or biases [4].

Recent studies have attempted to use similar features quantitatively – exploiting for instance HTML tags, URL patterns, number of affiliate links, and reuse of topic keywords – to estimate search engine optimization (SEO) and possible correlations with misinformation and page quality for specific domains such as product reviews [5], medical information [6].

Automatic approaches have the benefit of scalability, but inherently remain on the surface, making hard conclusions for specific guidelines impossible.

The topic of quality estimation is particularly relevant in the health domain, where false information is both wide-spread and dangerous, especially since the COVID-19 pandemic [7]. Related works manually analysed Google results to study specific phenomenons based on manually defined query terms, for instance for "immune boosting" [8, 9].

Apart from misinformation, factors like readability play a role when it comes to usefulness of information [10]. Related factors include fairness, accountability, transparency (FAT) [11], making the concept of relevance complex and subjective [12].

Large-scale analyses are difficult due the closed nature of query logs kept private by search engines. The AOL query logs historically provided an opportunity for quantitative query research including children queries [13], but their findings cannot be assumed to still apply today due to the ever-changing nature of search engine design and usage. Others use small-scale qualitative audience-specific research to, for instance, develop specific tools for improving of children queries [14].

---

[1]Data and full analyses: https://github.com/SlimZoeken/100queries

3. Data and Methodology

3.1. Data Collection

The essence of this research is a sample of authentic search queries, hence queries submitted by our target audience to an actual search engine. Using Google search, these queries lead to a set of search results which are the basis of our analysis.

3.1.1. Search Queries

Our query data originates from a random sample of queries submitted to a search engine for children (see Acknowledgements) that is used in schools. No personal data has been logged apart from the country and the ages of the children; they were in the Netherlands or Belgium and between 8 and 12 years old.

We have manually categorized the full raw sample of 200 queries into the following 11 categories (in order of frequency): Nature & Biology, Entertainment & Media, Society, noise (e.g. "haaaaaaaao"), Technology, History, Sports, Geography, General, Language & Culture, and Cooking.

We have removed the noise queries as well as and near-duplicates, e.g. spelling variations such as "Ronaldo" vs. "Cristiano Ronaldo"; while those variations might be interesting for search engine evaluation, they are not expected to yield insights into the slice of the web under investigation, which is the focus of this work.

Finally, we have down-sampled the queries proportionally per topic to reach a total sample size of 100 queries, including non-curricular examples like brand names as well as explicit sexual terms. We have chosen the number of 100 queries as a trade-off between sufficient sample diversity and representativeness on the one hand, and our available resources for manual in-depth research of the results on the other hand.

3.1.2. Search Results

In the next step, we have manually submitted each of these queries to Google, using a standardized configuration:

- Google Chrome browser
- Not logged in to Google (or elsewhere)
- All cookies accepted
- System language and preferred browser language set to Dutch
- No preferred language set in Google
- *Safe Search* not activated

We have manually logged all entries of the first results page shown by Google for each query, along with screenshots for potential error correction, validation and archival purposes. Basic data for each result page includes:

- the result URL
- the query that led to the result

- the rank on the Google result page
- the result source – the domain, or for larger platforms, for instance, a YouTube channel

Furthermore, we have logged the number of cookies blocked by the uBlock Origin browser plugin[2] in its default settings. This allows for analyses about potential correlation between the number of tracking cookies – as defined by said browser plugin – and page quality. In our data, we have not seen any significant correlation, however.

In most cases, the first Google search page shows 10 results, but this is not definite. We have encountered between 9 and 12 results on the first page. In total, we have logged 998 distinct results for the 100 input queries.

3.2. Result and Source Annotation

We have annotated each result separately on both the page and the source level. For individual result pages, we have defined the following categories with a fixed set of possible values:

- **Result type** corresponding to query intents [15]: `informative`, `transactional`, or `navigational`
- **Relevance** to the query: `yes`, `no`, or `maybe`
- **Readability** in relation to the target group: `complex`, `understandable`, `simple`, `too simple`

These categories give indications about the usefulness of a search result page for a specific user and their assumed intent, as expressed through a query. This part of the the analysis does not take a page's context into account and therefore mostly corresponds to traditional document-level search evaluation techniques [16].

Nevertheless, a page lives in a larger context, and is at least surrounded by the source which has published that page, and typically other pages as well as contextual information. In order to get more insights about intentions, credibility, trustworthiness, and underlying objectives playing in the landscape of interest, we have annotated each source by categories to contextualise the individual pages of a source.

In total, we have encountered 542 distinct sources for the 998 result pages introduced above, and annotated them in the following categories:

- **Primary purpose**; 32 distinct values, including: `product or company information`, `commercial information provider`, `knowledge base`, `web shop`, `government institution` etc.
- **Sector**; 31 distinct values, including: `offline business` (e.g shops and stores), `online-only business`, `government institution`, `publisher`, `NGO`, `independent knowledge base`, `travel & tourism`, `individual` (personal home-pages) etc.
- **Commercial intent** (based on manual site analysis): `not-for-profit` (no commercial intent), `commercial` (e.g. professional publishers, companies providing product info etc.), `ultra-commercial` (purely commercial sites without concern for content quality)

---

[2]uBlock Origin browser plugin: https://ublockorigin.com/

- **'About' page quality**: `comprehensive`, `sufficient`, `missing information` (incomplete), `contact info only`, `contact info in footer`, `no information`
- **Thumbs-up**; a curatorial judgement for whether a page is useful for the target group: `yes`, `just in case` (maybe), `no`, `unclear` (hard to judge)

There are inherent overlaps between some of these categories. For instance, the purpose of a business' web page is almost naturally to promote their own products, so most source in the 'offline business' sector will have a commercial intent and 'product or company information' as their primary purpose.

Furthermore, we have logged the owners of a site where possible. Turning this information into useful insights, however, requires further in-depth research to reveal potential sub-companies, straw men and nested organizational structures (see Section 6).

The annotations have been performed by two annotators, in a process developed over years in related settings in academic and applied works. The annotators reviewed, discussed and refined their annotations to converge into a single annotation per result.

4. Analysis and Results

4.1. The Queries

To get basic insights about the structure of our queries, we looked at the unfiltered sample of our 200 initial queries, and found:

- 74 queries comprise a **single word**
- 55 queries refer to a **proper noun** (*named entity*), e.g. a person or a place name.
- 19 nonsensical queries (**noise**)
- 16 queries are **questions**

4.2. The Search Results

Based on the 998 search result pages we received for our query sample, we investigated the 542 distinct sources and their distribution among those pages.

The most outstanding result of our data analysis is: Wikipedia is the one single source that dominates the search results across categories and dimensions, being by far the most frequent source across the result set, with Google-internal navigational links, e.g. to their own video search as runner-up. `wikikids.nl`, Facebook, and various encyclopedic or specialised pages follow in a Zipf-like shaped frequency distribution (Figure 1).

Wikipedia is even more dominant when looking at the highest ranked search result only. Google-internal links remain runner-up, being the only other source that hit the first rank more than once (Figure 2).

This allows two preliminary conclusions:

1. Most relevant information for our target audience is provided by Wikipedia.
2. There are no other individual sources that provide results that are broadly relevant for our target group.
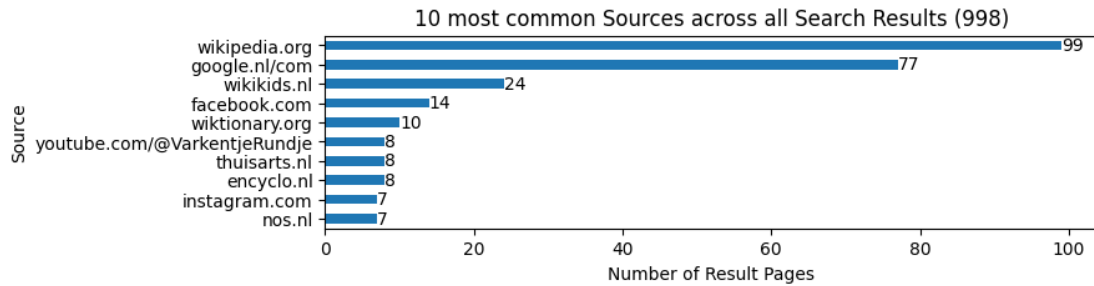
Figure 1: Among all results on the first Google result page, Wikipedia is the most frequent source. *google.nl/com* represents Google-internal links. The long tail of sources is cut off.
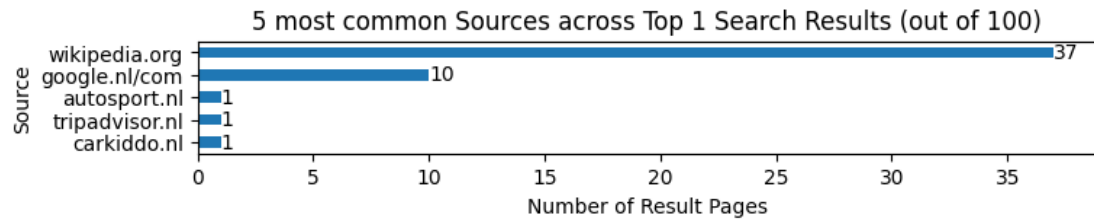
Figure 2: Looking at the highest-ranked search results only, Wikipedia is extremely dominant. The long tail is cut off.

Result Relevance For a search engine, topical relevance is considered one of the most signifiant benchmarks [17, 18, 19]. After all, the traditional goal of information retrieval is to identify documents that fulfil the information needs by a user as expressed by the search query, for which topical relevance forms a minimum requirement. As expected from the most established search engine, almost 90% of the results provided by Google have been topically relevant (Figure 3).
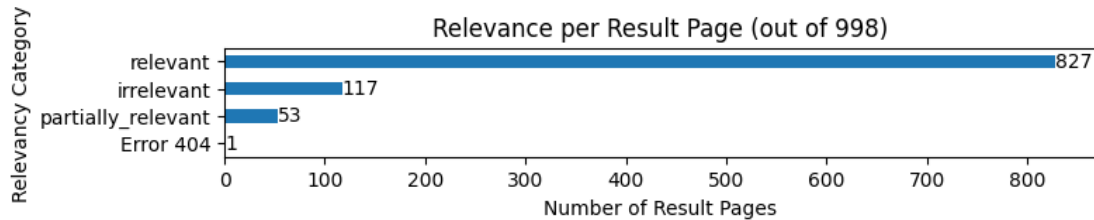

Figure 3: 88.2% of the Google results were indeed relevant for our target audience.

In the scope of this work, the topical relevance provided by search results primarily confirms that Google search serves as a mostly reliable guide into the region of the web that is relevant for the target audience as in: the page matches with the topic of a query – here, we are not looking into a broader relevance definition in the sense of usefulness, which can be subjective [20, 12].

Intent   The goal of our research, however, is not to evaluate search engine functionality, but to investigate the results as a representation for what our target group gets to see of the web, by means of using search results. As a first step towards understanding source motivations, we remain on the individual page level and analyse each result in correspondence to the query intent types. Following the established taxonomy of query intent [15], we have annotated which of these intents it addresses for each page: *informational*, *transactional* and *navigational*.

Whether or not the Google results align correctly with the respective query intents cannot be fully answered in all cases; ambiguity in the queries cannot always be resolved with the context available in our data. However, that question falls mostly into the field of search engine evaluation too. At the same time, we assume that the distribution of intent types we see in our data is representative in the part of the web that we investigate, so accuracy in individual cases is not our concern here.

Roughly 2/3 of the results are suitable for *informational* queries (Figure 4).
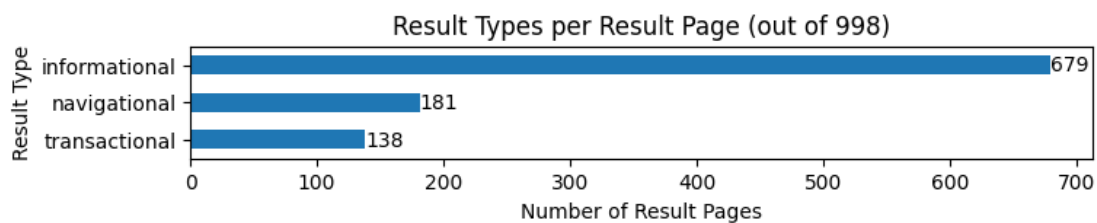

Figure 4: Query intents for which pages have been made primarily.

Advertisement   We have defined the amount of ads on a page as another potential indicator for commercial interest behind a source. However, when looking at a page, the plain number of ads does not seem the most significant factor. Other aspects such as their placement and their type play a crucial role to determine the perception of ads in relation to the page content. We have therefore defined five main level of advertisement presence on a page:

- *No ads*
- *Company promotion*: a company page that advertises its own products
- *Limited ads*: a page that contains ads, but the (informational) content dominates
- *Many ads*: a page in which ads are dominant, while the content is still readable
- *Over the top ads*: a page in which the actual content is hard to read, including for instance pop-up windows and similar distracting effects

Figure 5 shows their distribution among our search results:

Figure 6 shows how the number of ads increases the lower the rank of a result in the search results: looking only at the first ranks among the Google search results, almost 60% of the pages contain no ads at all – mostly thanks to the fact that Wikipedia is often found on the first rank. When taking all the results into account, however, that proportion of ad-free pages decreases to just over 30%; hence fewer than there are company promotion pages.
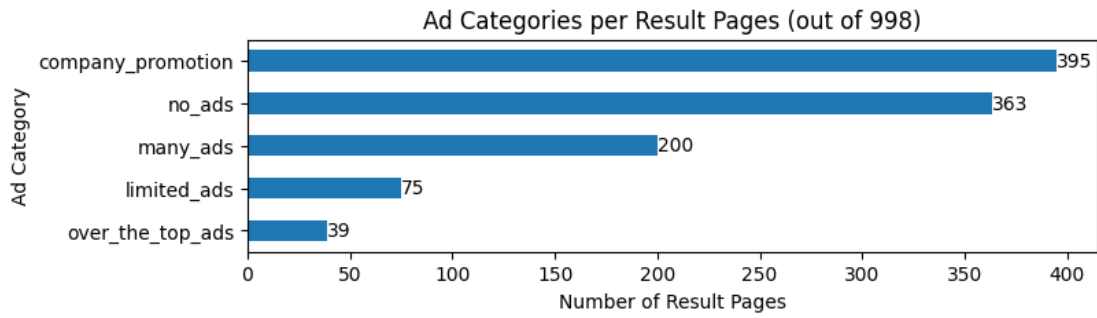
**Figure 5:** Advertisement level on our result pages.
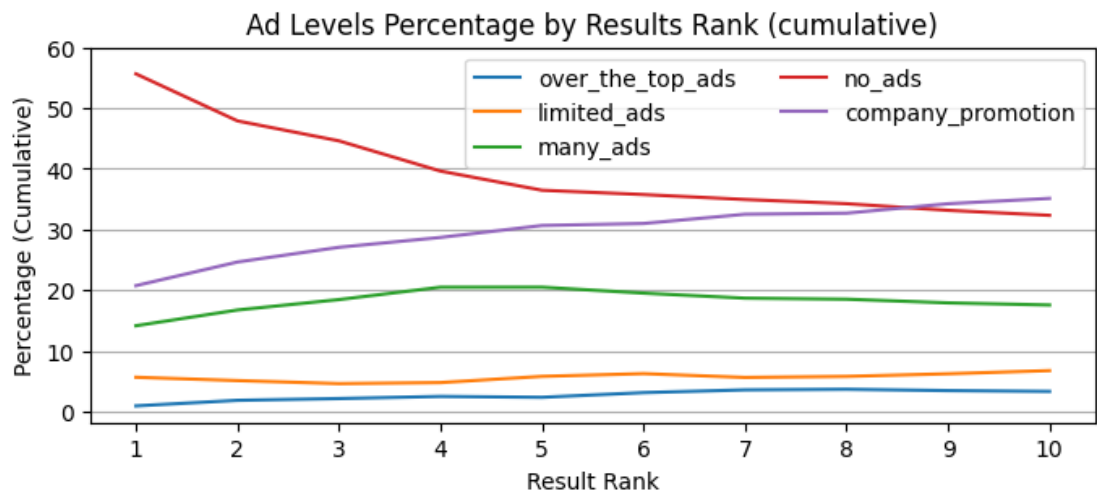
Figure 6: The proportion of results with *over the top ads*, cumulative up to rank N.

Thumbs-Up and 'About' Pages   In addition to the detailed annotations, our expert annotator gave a curatorial 'Thumbs-Up' annotation to each source. This is an informed, while subjective overall judgement about the general suitability for our target group. Figure 7 shows a clear correlation between pages that are ranked highly by Google and sources that we have evaluated positively. This indicates that the first few Google results are mostly useful in our use case, but also contain more than 10% of unsuitable sources.

In a related quest, we have manually investigated the 'About' pages or similar contact details provided by our sources, considering background information as a significant factor for accountability and transparency, and hence for credibility of a source [11]. 75% (753) or our search results provide sufficient or even extensive contact details.

Furthermore, we investigated the correlation between contact details and our judgement about source quality. We see that a large majority (76.8%) of our result pages from positively judged sources provide comprehensive contact details, while that portion goes down to 30.6% for 'Just in case' sources, and to merely 8.6% for sources that we consider as not suitable (Figure 8).
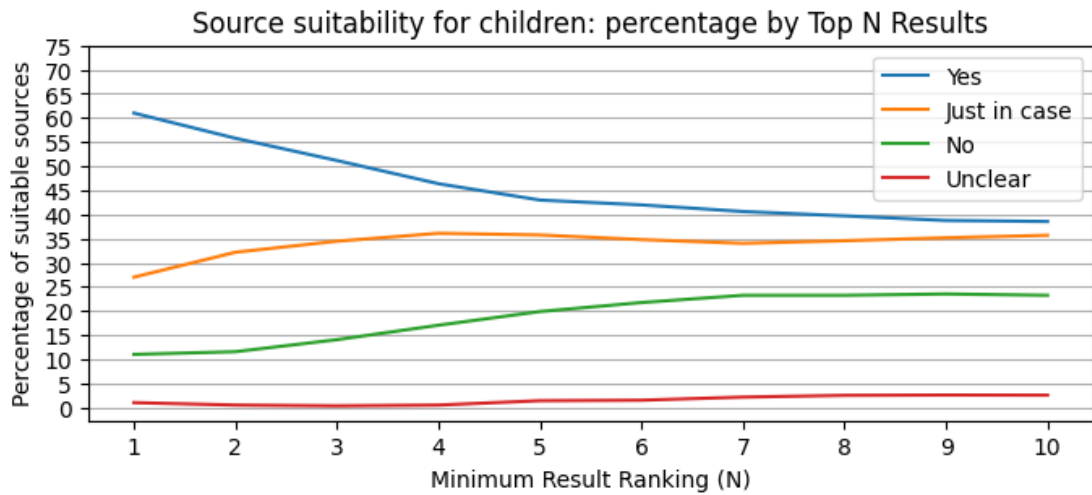
Figure 7: The proportion of good sources for our target group, cumulative up to rank N.

This finding makes the 'About' page a strong candidate for a significant factor for determining the quality of a source.
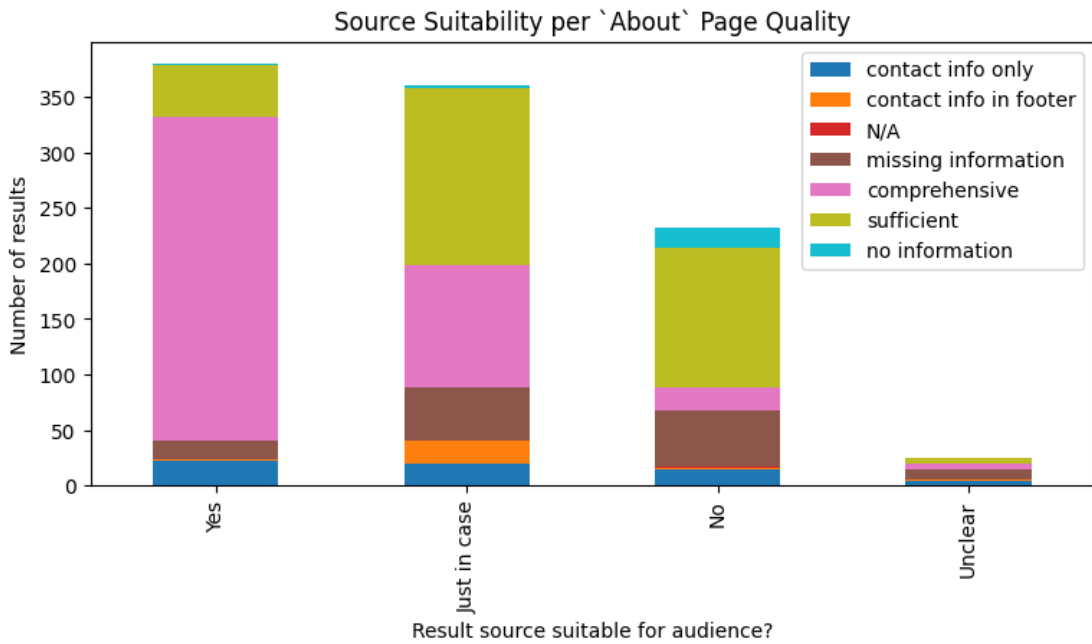
Figure 8: 'About' page and contact details correlate with source suitability judgements.

5. Conclusions and Discussion

5.1. Results

We have used a combination of authentic web queries and a manual evaluation of the respective Google search results to empirically get a grip on what the web looks like for a specific target group: Dutch children between 8 and 12 years old. From an educational perspective, the results can be summarized as: Wikipedia surfaces as primary source for information relevant to our target group. While it is generally of high quality in terms of correctness, Wikipedia is not designed specifically for children. No sources other than Wikipedia have been visible in our data consistently. Therefore, our target group navigates in a slice of web that is in principle accurate and trustworthy, but not necessarily suitable for our specific audience.

The relevance of non-commercial source has been expressed for other groups like breast cancer patients [10]. In our case, non-commercial sources have also proven to be most credible; apart from Wikipedia, this mostly refers to subsidised sources like *NOS* (part of the Dutch public broadcasting system) or domain-specific non-profit organizations like `thuisarts.nl`

Other frequently retrieved sources either fail to guarantee appropriate quality standards for different reasons, such as Facebook[3] and `wikikids.nl`[4], or are dictionaries which do not address the typically assumed information needs.

Another remarkable observation is that we did not encounter any sources that actively spread misinformation. This is presumably due to the non-political interests of our audience, as expressed in their search queries. It is still worth stating that the main issue for our audience is finding credible and relevant information, rather than identifying and filtering out misinformation.

This work is also designed to develop guidance to search practitioners. For our audience, one a set of practical tips in Dutch under CC BY-NC 4.0 license [5] [21]. This practical outcome welcomes translations into other languages and adaptations for other audiences.

5.2. Search Engine as a Lens

It is hard to definitely answer the question: "Is there enough content available on the web for a specific audience?" Anyway, we encourage practitioners to consider such questions before designing technical solutions to find such information. The method we present here takes a path to at least approximate such a question empirically.

With the presumption that Google search results for a query sample are representative for the relevant content that exists on the web, we have used internet search as a lens to get an empirical impression of slices of the web that are visible for a target audience. We have manually analysed the content found there to allow conclusions over the web and how members of a specific target group see it.

Using the metaphor of a lens to zoom into a specific portion of the web, this methodology can be applied for other target groups, for instance medical patients. For such purposes, we have developed a taxonomy that is generic enough to annotate pages from any other domain.

---

[3]On 7 January, 2025, Meta announced a stop on employing fact checkers for their platforms in the USA.
[4]*wikikids.nl* is written for and by children only with informal help of adults. There is no systematic quality control.
[5]Practical guidance regarding children internet literacy (in Dutch): https://slimzoeken.nu/online

Designing and refining this taxonomy has been an interactive process. We see its current state as a subset for all possible categories and values that covers the requirements for our use case. We consider it to be transferable to other applications, which might, however, require extensions to our subset. For instance, as we mentioned, active misinformation is not an issue in our case, but it might be in other fields, including the health domain – requiring respective annotations of case-dependent granularity.

Limitations   We use Google as a lens to look at a slice of the web that is relevant for a specific group. However, Google Search is a hardly customisable product designed for a different purpose, and assumptions made over query intentions, user preferences and sources are not transparent. Therefore, our data only reflects a snapshot that is a result of specific states of the search index and the search algorithm, both of which change continuously.

The query sample that we use originates from a single search engine, more research based on a query with more diverse provenances such as different search engines, user interfaces, and contexts. Generally, we can only speculate about reasons on why we see certain results while others remain invisible. For instance, our annotator noticed a remarkable absence of well-known sources with high relevance for the target group, including the public broadcasting programmes "schooltv.nl", "Jeugdjournaal" (*Youth News*), and "Klokhuis". Out of these, the latter two are not retrieved at all, and "schooltv.nl" occurs just three times in the form of outdated PDF files.

These cases can likely be explained with Google's attempt to separate video results from the general results. An ad-hoc investigation in the Google Video search with the same query sample shows that these sources are much more present there. Other sources might be affected in different ways by internal search engine logic beyond our knowledge.

6. Future Work

The main directions of potential follow-ups to this research concern refining the presented methodology for web research, and applying it in other contexts.

6.1. Methodological Refinement

Query Collection Challenges   The main challenge for applying our methodology is getting a sample of authentic queries that is representative for a target audience. Search engine providers have access to raw query data, but most commercial providers are not willing to share or to publish that valuable part of their intellectual property – even though a query could be seen as personal data of the person that submitted it.

Technological approaches of logging queries independently of the search engine, for instance on the client-side are at risk to introduce a sample bias towards users who actively contribute to such approaches, for instance by installing a browser extension for that purpose.

However, we hope that other domain-specific or site-internal search engines – for instance integrated in consumer-facing medical sites – could yield search query samples that approximate the general information needs by their specific target group. Hopefully, such sites are more

open to cooperate with researchers due to their background being either non-commercial or focussing on products that are not search-related.

Another approach might be synthetically generated queries, for instance using generative language models. Doubts about their representativeness and hence the results of a study based on such queries, however, can currently not be resolved.

Underlying Networks   In the present study, we investigate sources individually to provide context to individual search results. We we have seen indications for connections on another level when looking at the owners of those pages.

In following iterations of this research, we are planning to systematically log the source owners, too. These information should give even more context about site owners and their underlying connections.

To come to full use, the parent organizations behind page owners should be part of such research, too, as far as applicable. This also increases the effort, but seems like an important next step when it comes to transparency and credibility.

Chatbots and Generative AI   For multiple reasons, we are confident that search engine-based web research will remain highly relevant even while facing the raise of generative language models in the recent years: Chatbots driven be 'Large Language Models' (LLMs) are not suitable for information access and retrieval on their own [22]. Approaches such as RAG (*retrieval-augmented generation*) [23] still rely on search engines for making generative models more accurate and relevant.

Chatbots have been proposed as means for information access, also in an educational context[6]. Anyway, the question of whether there is sufficient coverage that can feed language model-based approaches for a domain thus remains crucial.

Furthermore, the method of using authentic input can be transferred to evaluating chatbots using prompts instead of queries [24], even though scientific evaluation of systems that do not produce reproducible outputs face additional challenges.

In principle, we consider the presented methodology as mostly agnostic to future methods of information access. Query samples will change in the light of new trends and topics of interest, but can be evaluated in the same way.

6.2. Other Applications

We are looking into applying the present method in other domains, for instance the afore-mentioned medical domain. Another suggestion is to use the same methodology with a different query sample from the same user group in order track changes over time. Similarly, we are planning to use search engines other than Google for both validation and comparison of the results.

---

[6]ChatGPT Edu: https://openai.com/index/introducing-chatgpt-edu/

6.3. What is Quality?

Perhaps more interestingly from a practical point of view, our data analysis combined with human expertise provides empirical insights about what makes a good source. Features that seem unrelated to the quality of a page at first glance – for instance existence and comprehensiveness of an 'About' page, the number of tracking cookies, moderation mechanisms etc. – could be used to develop semi-automatic systems that estimate the quality as well as the intention of an unseen source.

Classical network algorithms such as HITS [25] and PageRank [26] for authority identification and ranking respectively, or for community detection [27], could be applied to identify clusters of (non-)credible communities within a slice of the web.

For practical reasons, the quality of a page has often been seen in a decontextualized manner: a page contains either good or bad information. This view results from and impacts the tasks expected from a search engine: identify bad pages, and rank the relevant ones.

Especially, but not exclusively when it comes to education, however, the quality of a page is much more complex. While we have not encountered active misinformation, we have identified sources that have good intentions, but lack educational concepts or effective mechanisms to ensure high quality of their pages. In particular, community-driven sites such as Wikis face the risk of becoming sources for useless and misleading pages if they lack domain-specific knowledge or sufficient resources for moderation and quality control.

In future work, narrowing down the definition of quality in a feedback loop with semi-automatic quality estimates will be another line of research.

Acknowledgments

Declaration on Generative AI

The authors have not employed any Generative AI tools.

References

[1] S. Nakamura, S. Konishi, A. Jatowt, H. Ohshima, H. Kondo, T. Tezuka, S. Oyama, K. Tanaka, Trustworthiness Analysis of Web Search Results, in: L. Kovács, N. Fuhr, C. Meghini (Eds.), Research and Advanced Technology for Digital Libraries, Springer, Berlin, Heidelberg, 2007, pp. 38–49. doi:10.1007/978-3-540-74851-9_4.

[2] Y. Yamamoto, K. Tanaka, Enhancing credibility judgment of web search results, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI

---

[7]WizeNoze: https://www.wizenoze.com/

'11, Association for Computing Machinery, New York, NY, USA, 2011, pp. 1235–1244. doi:10.1145/1978942.1979126.

[3] R. Ennals, B. Trushkowsky, J. M. Agosta, Highlighting disputed claims on the web, in: Proceedings of the 19th International Conference on World Wide Web, WWW '10, Association for Computing Machinery, New York, NY, USA, 2010, p. 341–350. doi:10.1145/1772690.1772726.

[4] R. White, Beliefs and biases in web search, in: Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval, SIGIR '13, Association for Computing Machinery, New York, NY, USA, 2013, pp. 3–12. doi:10.1145/2484028.2484053.

[5] J. Bevendorff, M. Wiegmann, M. Potthast, B. Stein, Is Google Getting Worse? A Longitudinal Investigation of SEO Spam in Search Engines, in: N. Goharian, N. Tonellotto, Y. He, A. Lipani, G. McDonald, C. Macdonald, I. Ounis (Eds.), Advances in Information Retrieval, volume 14610, Springer Nature Switzerland, Cham, 2024, pp. 56–71. doi:10.1007/978-3-031-56063-7_4.

[6] S. Schultheiß, H. Häußler, D. Lewandowski, Does Search Engine Optimization come along with high-quality content? A comparison between optimized and non-optimized health-related web pages, in: Proceedings of the 2022 Conference on Human Information Interaction and Retrieval, CHIIR '22, Association for Computing Machinery, New York, NY, USA, 2022, pp. 123–134. doi:10.1145/3498366.3505811.

[7] I. A. Portillo, C. V. Johnson, S. Y. Johnson, Quality Evaluation of Consumer Health Information Websites Found on Google Using DISCERN, CRAAP, and HONcode, Medical Reference Services Quarterly 40 (2021) 396–407. doi:10.1080/02763869.2021.1987799.

[8] A. Cassa Macedo, A. Oliveira Vilela de Faria, P. Ghezzi, Boosting the Immune System, From Science to Myth: Analysis the Infosphere With Google, Frontiers in Medicine 6 (2019). doi:10.3389/fmed.2019.00165.

[9] C. Rachul, A. R. Marcon, B. Collins, T. Caulfield, COVID-19 and 'immune boosting' on the internet: a content analysis of Google search results, BMJ Open 10 (2020) e040989. doi:10.1136/bmjopen-2020-040989.

[10] Y. Li, X. Zhou, Y. Zhou, F. Mao, S. Shen, Y. Lin, X. Zhang, T.-H. Chang, Q. Sun, Evaluation of the quality and readability of online information about breast cancer in China, Patient Education and Counseling 104 (2021) 858–864. doi:10.1016/j.pec.2020.09.012.

[11] D. Shin, Y. J. Park, Role of fairness, accountability, and transparency in algorithmic affordance, Computers in Human Behavior 98 (2019) 277–284. doi:10.1016/j.chb.2019.04.019.

[12] C. Shah, E. M. Bender, Envisioning Information Access Systems: What Makes for Good Tools and a Healthy Web?, ACM Trans. Web 18 (2024) 33:1–33:24. doi:10.1145/3649468.

[13] S. Duarte Torres, D. Hiemstra, P. Serdyukov, An analysis of queries intended to search information for children, in: Proceedings of the third symposium on Information interaction in context, IIiX '10, Association for Computing Machinery, New York, NY, USA, 2010, pp. 235–244. doi:10.1145/1840784.1840819.

[14] J. A. Fails, M. S. Pera, O. Anuyah, C. Kennington, K. L. Wright, W. Bigirimana, Query Formulation Assistance for Kids: What is Available, When to Help & What Kids Want, in: Proceedings of the 18th ACM International Conference on Interaction Design and

Children, ACM, Boise ID USA, 2019, pp. 109–120. doi:`10.1145/3311927.3323131`.

[15] A. Broder, A taxonomy of web search, SIGIR Forum 36 (2002) 3–10. doi:`10.1145/792550.792552`.

[16] S. Buttcher, C. L. A. Clarke, G. V. Cormack, Information Retrieval: Implementing and Evaluating Search Engines, Mit Pr, Cambridge, Massachusetts London, England, 2010. URL: https://plg.uwaterloo.ca/~ir/ir/book/.

[17] S. Mizzaro, Relevance: The whole history, Journal of the American Society for Information Science 48 (1997) 810–832. doi:`10.1002/(SICI)1097-4571(199709)48:9<810::AID-ASI6>3.0.CO;2-U`.

[18] P. Borlund, The concept of relevance in IR, Journal of the American Society for Information Science and Technology 54 (2003) 913–925. doi:`10.1002/asi.10286`.

[19] B. Hjørland, The foundation of the concept of relevance, Journal of the American Society for Information Science and Technology 61 (2010) 217–237. doi:`10.1002/asi.21261`.

[20] T. Saracevic, Relevance: A review of the literature and a framework for thinking on the notion in information science. Part II: nature and manifestations of relevance, Journal of the American Society for Information Science and Technology 58 (2007) 1915–1933. doi:`10.1002/asi.20682`.

[21] M. Sprenger, C. Schnober, Zoeken onderzocht – Het 100 Queries project, Taalunie HSN-archief (2024). URL: https://hsnbundels.taalunie.org/bijdrage/zoeken-onderzocht-het-100-queries-project/.

[22] C. Shah, E. M. Bender, Situating Search, in: Proceedings of the 2022 Conference on Human Information Interaction and Retrieval, CHIIR '22, Association for Computing Machinery, New York, NY, USA, 2022, pp. 221–232. doi:`10.1145/3498366.3505816`.

[23] P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, H. Küttler, M. Lewis, W.-t. Yih, T. Rocktäschel, S. Riedel, D. Kiela, Retrieval-augmented generation for knowledge-intensive NLP tasks, in: Proceedings of the 34th International Conference on Neural Information Processing Systems, NIPS '20, Curran Associates Inc., Red Hook, NY, USA, 2020, pp. 9459–9474. URL: https://proceedings.neurips.cc/paper_files/paper/2020/file/6b493230205f780e1bc26945df7481e5-Paper.pdf.

[24] R. Muehlhoff, M. Henningsen, Chatbots im Schulunterricht: Wir testen das Fobizz-Tool zur automatischen Bewertung von Hausaufgaben, 2024. doi:`10.48550/arXiv.2412.06651`.

[25] J. M. Kleinberg, Authoritative sources in a hyperlinked environment, J. ACM 46 (1999) 604–632. doi:`10.1145/324133.324140`.

[26] L. Page, Method for node ranking in a linked database, 2006. URL: https://patents.google.com/patent/US7058628B1/en.

[27] M. Girvan, M. E. J. Newman, Community structure in social and biological networks, Proceedings of the National Academy of Sciences of the United States of America 99 (2002) 7821–7826. doi:`10.1073/pnas.122653799`.