From Alibaba to YouTube: User Searches for Digital Democracy Topics

Christoph Glauser[1] and Uwe Serdült[2,3]

[1]Institute for Applied Argumentation Research IFAAR, Berne, 3011, Switzerland
[2]Center for Democracy Studies Aarau (ZDA), Aarau, 5000, Switzerland
[3]Ritsumeikan University, Kusatsu, 525-8577, Japan

ABSTRACT

Digital democracy tools such as e-consultations, e-petitions or internet voting are playing an increasing role in the public debate. They have become an important part of the digitization process in politics and governments. A digital lifestyle in general and during the pandemic has led to a higher demand for so-called civic tech tools. For this reason, during the Corona Summer 2021, search terms related to digital democracy were monitored for several months across various digital channels in Switzerland. Search term behavior was studied in German, French and Italian as the three main languages spoken in Switzerland. The results show that there is a significantly higher active search volume for e-democracy and e-participation topics in the German-speaking region than in the French- and Italian-speaking regions. Internet users search for such topics in a great variety of different international and local search engines as well as on social media channels that go beyond only Google or Facebook.

Keywords: Digital democracy, Digital votes, Online voting, e-democracy, e-voting, e-elections, e-participation, Search behavior, Local search

INTRODUCTION

The Corona crisis has had a rather negative impact on democracy in many countries (Boese et al. 2022; Engler et al. 2021). First research in the field shows that the level of democracy has been hampered during the pandemic, whereas the public interest on online channels and social media networks for democratic topics on the other hand has significantly increased. It is therefore crucial to further examine the perception of the population to better evaluate the user-, or citizen centricity during such a global crisis. Therefore, the online search behavior of the Swiss online population during the COVID-19 pandemic and during the application of the stepwise exit plan of the Swiss Government during summer 2021, between two public votes and before the third wave of October 2021 is analyzed. The authors monitor keywords of different categories of digital democracy topics such as for example "online Wählen", "digitale Initiative", "participation politique", "élection en ligne", "impegno online" or "identità digitale".

Beyond data science itself, some other fields in the social sciences or medicine do also benefit from the power of *big data and AI*. This paper illustrates a specific use case on how such approaches have been applied in the context of

the online communication on e-democracy and e-participation related issues and topics. This use case illustrates how data science and big data can be used as a powerful instrument to monitor the amplitudes of the searches on digital democracy.

The basic research questions of the project are the following: a) Which ones and to what extent are digital democracy topics already known and most frequently used by Swiss users and citizens. b) What are the topics related to democracy in their generic digital communication use and c) on which of their favorite online search and social media channels do Swiss citizens search most frequently for new digital democracy topics? Our work continues to explore the feed-back mechanisms of citizens, to better understand the public debate on the future of democracy.

Previous work has shown that online searches for COVID-19 are strongly correlated with international WHO announcements (Szmuda et al. 2020). These studies provide interesting insights; however, they do each only cover a rather small percentage of the population. In the present study the entire online population of Switzerland, i.e., 8.1 million active internet users, is covered by analyzing the search behavior on a large number of online channels. A recent study has investigated the information and communication behavior of the German speaking Swiss population during the first phase of the Corona crisis (Friemel et al. 2020). The latest empirical results of this monitoring are presented and discussed.

DATA AND METHODS

The methods applied in this monitoring are based on presenting an online science communication framework with a dictionary of keywords which semantically describe the most recent public discussion on e-democracy, e-participation and e-voting or e-collection of votes in Switzerland. The framework explores on one hand previous research topics on digital democracy (Huang et al. 2021; Mendez and Serdült, 2017; Serdült et al. 2016) and on the other hand on existing dictionaries from the political science literature in the field of e-democracy (Hennen et al. 2020; OECD, 2020; UNDESA, 2020). It led first to new pathways of measuring digital communication systematically by API tracking of what users are searching already on COVID topics by using multiple digital platforms in Switzerland (Glauser et al. 2020). Then the authors decided to apply the approach to extensive digital democracy issues for a continuous monitoring.

The method relates to comparisons between the digital content that is offered on e-democracy and e-participation and the related relevant content, which is searched for in a defined online universe (mostly based on a country, or on different country domains and where these are missing, based on ITU[1] user data) working with comparable data from as many different sources as possible. This approach is unique because it does not include secondary media information, but rather the active searches by the user community

---

[1]International Telecommunication Union (ITU). Internet live stats. Website, 2016. Available https://www.internetlivestats.com/internet-users-by-country/; checked June 24th, 2020.

itself, measured by API or software interfaces on the respective platforms where any kind of search field is offered to the user. All API data and search data, which is directly provided by the suppliers, is thoroughly parametrized by several different parameters like timeframe, currency, keyword frequency, keyword trends, technology, language etc. In comparison to classical surveys this approach has the advantage that there is no interference between the researcher or the questionnaire and the interviewee.

The search behavior of the online population has since 2000 constantly improved and has become much more differentiated than it was in the beginning of online usage. Today, depending on the target group, "special interest" keywords and combinations of them are also often searched on platforms such as LinkedIn, Wikipedia, Facebook, or Instagram. In addition to this, as the authors' experience shows, on average 30-40% of the universal search today is happening on social media, e-shops and is basically also mobile search. New search options are also constantly being added while old ones are disappearing. These changes are becoming increasingly important to online research. The power of search data has been proven in several use case studies to measure the attitude, the sentiment, or the opinion of the broad population, for example to study racial animus during the elections (Stephens-Davidowitz, 2014) or to detect feelings such as low self-esteem (Zaman et al. 2019). It has been shown that such valid and reliable methods can replace the results of traditional surveys, since participants tend to justify and rationalize their own behavior in surveys (Stephens-Davidowitz and Pabon, 2017). Furthermore, it has been shown that only a specific part of the population participates in surveys which can lead to biased results (Wright, 2005). It has been estimated that 87.2% of the Swiss population[2] are using the internet and thus different types of search engines, social media, and e-shops.

Most contemporary research on search is often limited to Google data (Wilde et al. 2020) or on social media like Twitter (Senaka et al. 2020) hashtags with the strong limitation that only one part of the active internet population is covered. This study uses a new cross channel approach, extracting data from 14'103 sources worldwide including search engines (e.g., Wikipedia, Google, Bing, Yahoo!, Ask, Lycos, Alexa, Technorati, MetaCrawler, Search.com, local SE's etc.), social network platforms (e.g., Facebook, Twitter, LinkedIn etc.) and e-shop searches (e.g., Amazon, Ebay, Alibaba etc.). The harvested data is analyzed systematically based on the parameters of each channel to reveal the queries submitted by internet users and the frequencies of those queries by the domains of 203 countries. The possibility to examine, parametrize and compare numerous search engines and social networks allows to gain valid and reliable aggregated horizontal insights for the first time into the internet search behavior in terms of content of the digitally active Swiss population, counting approximately 8.1 million users in Switzerland. Using this multi-channel approach, some of the challenges in

---

[2]International Telecommunication Union (ITU). Internet live stats. Website, 2016. Available https://www.internetlivestats.com/internet-users-by-country/; checked June 24th, 2020.

Proximity and Distance of Search Demand:
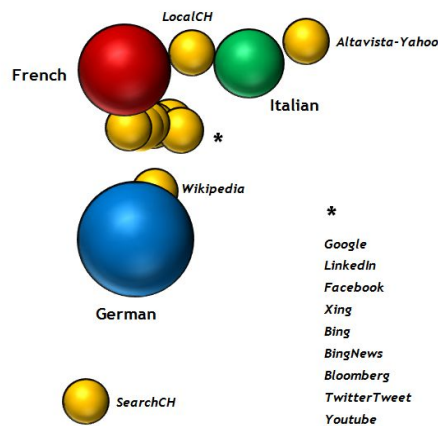Channels and Languages in Switzerland

Figure 1: Proximity and distance of the languages, klein geschrieben and channels over the observed time frame. Search platforms and language searches show proximities and distances in 3D.

big data (as shown for example Khan et al. 2019), such as variability and variety, can be overcome.

The crawl is adjusted to consider only a predefined geographical country based on the top-level internet domain. In this study the focus was put on Switzerland. The scope of this work was defined on e-democracy and e-participation keywords in German, French, Italian and English[3] to target the Swiss population and its active searches. The digital monitoring period started on April 5th and ended on September 30th, 2021.

RESULTS

Figure 1 depicts the proximity and distance between the different channels and language regions in Switzerland between April 5th and September 30th. It shows in three dimensions without scale in 3D, which channels are used in which languages most frequently when it comes to e-democracy topics and issues. The Latin languages French and Italian searches are much closer to each other than the German which is closer to French than to Italian user demand. The size of the language bubbles shows the number of searches in the three languages. The size of the bubble represents quite well the language regions of Switzerland - Swiss German population is on average 62.3% of the entire population, French speaking is 22.8% and Italian is at 8%. The most prominent channels are used equally in all the languages. The distributed usage of channels like Google, Facebook, Instagram, LinkedIn, Xing, Bing, Bing News, Twitter, gross geschrieben, etc. is illustrated by the accumulation

---

[3]The English semantic dictionary was also measured and only used as a control group of keywords. But the English keyword search was not used for the interlingual comparisons, since the official Swiss languages are German, French, and Italian.
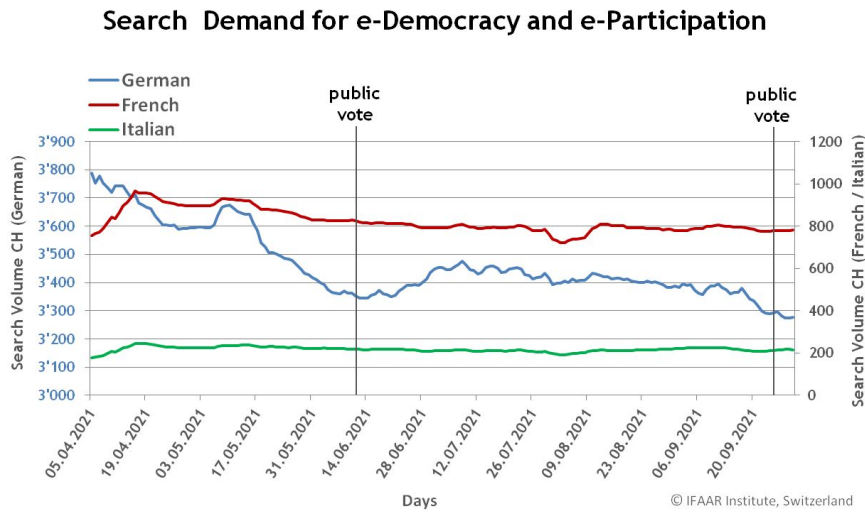
Search Demand for e-Democracy and e-Participation

Figure 2: Search behavior on e-democracy and e-participation keywords: The respective average of all keywords in German, French and Italian is depicted. Each data point shows the average search volume over all channels for the last 30 days in Switzerland.

of yellow "bubbles" in the center of the 3D figure and show that the digital democracy topics are discussed on most of the relevant online platforms in Switzerland.

Figure 2 shows the digital search behavior of the Swiss online population for the different language regions. The analysis included search volume measurements during the time of analysis. Each data point shows the daily moving average search volume over all the channels for the previous 30 days. The search volume results of all the keywords measured which have been aggregated and Figure 2 shows the compared search volume for all German/French/Italian keywords. The figure illustrates that the interest in e-democracy keywords was high by the German and French speaking Swiss users always in the weeks before a public vote. Whereas the French and Italian keywords are remaining on an ongoing constantly lower level.

The table 1 shows the average search volume per channel over 30 days of the most searched topics in the three languages analyzed (German, French, Italian). Whilst the German speaking population is mainly interested in practical topics like online voting, online election ("online Wählen", "online Stimmen", "digital Wählen", "digital Stimmen") or digital initiative, the French and Italian speaking Swiss population also shows interest in topics like engagement and political participation ("participation politique", "engagement en ligne", participation en ligne", "impegno online", "partecipazione politica"). Furthermore, the interest in the often discussed digital identity ("identité en ligne", "identità digitale") is more present in French and Italian, compared to German. Overall, the keywords with the highest search demand are the ones concerning the ability to participate online by voting and electing electronically.

Table 1. Average search volume per channel over 30 days of the top 5 keywords per language during the analysis period.

Average Search Volume: Top 5 Keywords per Language

German		French		Italian	
Keyword	Search Volume	Keyword	Search Volume	Keyword	Search Volume
online Wählen	36'568	engagement en ligne	4'836	impegno online	1'269
online Stimmen	17'045	participation politique	3'151	identità digitale	1'050
digital Wählen	10'484	identité en ligne	2'917	iniziativa online	908
digital Stimmen	7'467	élection en ligne	2'790	identità online	881
digitale Initiative	4'134	participation en ligne	2'622	partecipazione politica	530

CONCLUSION

On a horizontal level the data show in Figure 2 that in the French-and Italian speaking regions of Switzerland the topic of digital democracy is a rather stable topic. Whereas in the German part the data show some significant fluctuation of the search activities in relation to the public votes. Before public votes the discussion about e-democracy and e-participation topics is increasing and shortly before the actual voting date, the search demand is decreasing again. This shows that the topic is volatile. However, there are also a few very interesting findings shown on Figure 1 looking at Wikipedia, SearchCH, LocalCH and Altavista-Yahoo. Whereas SearchCH as a rather German local search engine, is mainly used in the German speaking part of Switzerland, obviously Altavista-Yahoo, who's Yahoo headquarters have long been near Lausanne, is still more prominently used by the Italian and French speaking users. LocalCH usage is situated between French and Italian while Wikipedia is also much more intensively used in the German language for the digital democracy topics. All the other search possibilities are situated in the middle of the three language bubbles, meaning that they are equally used by all Swiss users together in more or less the same number of active searches.

ACKNOWLEDGMENT

REFERENCES

Boese, V.A., Alizada, N., Lundstedt, M., Morrison, K., Natsika, N., Sato, Y., Tai, H., and Lindberg, S.I. (2022) Autocratization Changing Nature? Democracy Report 2022. Varieties of Democracy Institute (V-Dem).

Engler, S., Brunner, P., Loviat, R., Abou-Chadi, T., Leemann, L., Glaser, A. and Kübler, D. (2021) "Democracy in times of the pandemic: explaining the variation of COVID-19 policies across European democracies", in: West European Politics, Volume 44 No. 5-6, pp. 1077-1102.

Friemel, T.N., Geber, S. and, Egli, S. (2020) Informations- und Kommunikationsverhalten in der Coronakrise. Technical report, Institut für Kommunikationswissenschaft und Medienforschung (IKMZ), University of Zurich.

Glauser, C., Schmid, L., and Kurpicz-Briki, M. (2020) 'Identifying the Public Interest in COVID-19 Contact Tracing Apps in Switzerland Based on Online Search Behavior', Proceedings of the Thirteenth International Conference on Theory and Practice of Electronic Governance, Athens, Greece.

Hennen, L., van Keulen, I., Korthagen, I., Aichholzer, G., Lindner, R., Nielsen, R.O. ed. (2020). European E-Democracy in Practice. Cham: Springer.

Huang, H.-Y., Kovacs, M.; Kryssanov, V.; Serdült, U. (2021). Towards a Model of Online Petition Signing Dynamics on the Join Platform in Taiwan, ICEDEG2021 – 8th International Conference on eDemocracy & eGovernment, Quito, Ecuador, 28–30 July 2021, pp. 199–204.

Khan, N., Naim, A., Hussain, M. R., Naveed, Q. N., Ahmad, N., and Qamar, S. (2019, May). The 51 v's of big data: survey, technologies, characteristics, opportunities, issues and challenges. In Proceedings of the international conference on omni-layer intelligent systems, pp. 19–24.

Mendez, F. and Serdült, U. (2017) What drives fidelity to internet voting? Evidence from the roll-out of internet voting in Switzerland, Government Information Quarterly, 34 (3), p. 511–523. doi: 10.1016/j.giq.2017.05.005

OECD (2020) Innovative Citizen Participation and New Democratic Institutions: Catching the Deliberative Wave: Highlights 2020. Paris, OECD.

Serdült, U.; Mendez, F.; Harris, M. and Seo, H.S. (2016) Scaling Up Democracy with E-Collection? CeDem 2016 Conference for E-Democracy and Open Government 2015, 18–20 May 2016, Danube University Krems, Austria, pp. 25–31.

Senaka F., Díaz López, J.A., Şerbana, O., Gómez-Romero, J.; Molina-Solana, M., and Guoa, Y. (2020) Towards a large-scale twitter observatory for political events, Future Generation Computer Systems, 110, pp. 976–983.

Stephens-Davidowitz, S. (2014) The cost of racial animus on a black candidate: Evidence using google search data. Journal of Public Economics, 118, pp. 26–40.

Stephens-Davidowitz, S. and Pabon, A. (2017) Everybody lies: Big data, new data, and what the internet can tell us about who we really are. HarperCollins New York.

Szmuda, T., Ali, S., Hetzger, T.V., Rosvall, P., and Sloniewski, P. (2020) Are online searches for the novel coronavirus (covid-19) related to media or epidemiology? A cross-sectional study. International Journal of Infectious Diseases.

United Nations, Department of Economic and Social Affairs - UNDESA (2020) United Nations E-Government Survey 2020: Digital Governance in the Decade of Action for Sustainable Development. New York, United Nations.

Wilde, J., Chen, W., and Lohmann, S. (2020): COVID-19 and the Future of US Fertility: What Can We Learn from Google?; IZA – Institute of Labor Economics, Germany.

Wright, K.B. (2005) Researching internet-based populations: Advantages and disadvantages of online survey research, online questionnaire authoring software packages, and web survey services. Journal of computer-mediated communication, 1034, 10 (3).

Zaman, A., Acharyya, R., Kautz, H., and Silenzio V. (2019) Detecting low self-esteem in youths from web search data. In The World Wide Web Conference, pp. 2270–2280.