SEARCH ATLAS

Visualizing Divergent Search Results Across Geopolitical Borders

Rodrigo Ochigame*
Massachusetts Institute of Technology
Cambridge, MA, USA
ochigame@mit.edu

Katherine Ye*
Carnegie Mellon University
Pittsburgh, PA, USA
kqy@cs.cmu.edu
equal contribution

ABSTRACT

The internet is full of invisible borders—geographic, linguistic, cultural, political—that circumscribe the information each user sees. Search engines shape such "information borders" by tailoring results according to geolocation, language, and other user profiling. We present Search Atlas, a tool paired with visualizations that enable users to see and cross these borders. For instance, how do search results for the same query differ for Brazilian, Turkish, and Indian users? Given a query, the tool displays multiple lists of Google search results, highlighting distinctive words for each set of parameters. Then, we provide visualizations that juxtapose and cluster Google results across countries, revealing new information borders and regions that can vary widely depending on the query. By exposing the partial perspective of a search engine, Search Atlas invites users to experience the internet from divergent positions and to reflect on how their online lives are conditioned by technological infrastructures and geopolitical regimes.

Fig. 1. A world map of "god" according to Google. This map shows top image results for translations of "god" in the default languages of most Google-supported countries, overlaid on the approximate geographic location of each country. Christian-majority countries tend to show images of a Christian god, Muslim-majority countries tend to show images of Allah (in written Arabic), and Buddhist-majority countries tend to show images of Buddha.

INTRODUCTION

How many times did you Google something today? Think of all the search results you have seen over the years, and how those results have gradually shaped your opinions, behaviors, identities, worries, and hopes; your ideology, your friend circle, and your worldview.

Trusting Google is understandable. Google tries to create the impression of a benevolent, all-seeing god. It claims that its search algorithms "sort through hundreds of billions of webpages" in an index that "contains more info than in all the world's libraries put together" in order to "find the most relevant, useful results for what you're looking for" [1]. This impression is only reinforced by the minimal design of Google's search engine interface, which omits its "partial perspective" [2], that is, the combination of choices that inevitably exclude some points of view in favor of others. Fig. 2 illustrates some of the invisible processes that shape the production of search results.

Search engine design is not just a technical matter, but a political one. Designers make consequential political choices regarding which sites to include and which to exclude, how to rank the included sites, and how to determine a site's "relevance" for a given query. For more than two decades, an expansive body of research has queried the politics of search engines. Even the earliest studies, based on anecdotal observations, already suggested that search engines systematically suppress some sites in favor of others, in line with financial interests [3]. More recent studies have argued that commercial search engines deploy algorithms that reinforce existing social structures, particularly racist and sexist patterns of exposure, invisibility, and marginalization [4]. Thus, it is vital to expose the partial perspective of search engines.

Yet, researchers face a recurring challenge: since the algorithms of commercial search engines are proprietary and secret, it is difficult to gather empirical evidence about their social effects [5, 6, 7]. But even in the absence of access to proprietary algorithms, it is possible to study their outputs: search results. Some studies have documented differential patterns in search engine indexing, for example finding that U.S.-based sites were more likely to be indexed by major search engines than their counterparts based in other countries such as China [8]. Other studies have compared search results for different queries, for example showing that Google searches for first names associated with Black Americans were more likely to yield discriminatory ads that suggested arrest records [9].

An even greater challenge is to study how search results for the *same* query differ for different users. Many search engines tailor results according to geolocation, language, and other user profiling. In this sense, the internet is full of "information borders" that users cannot easily cross. And despite an abundance of public discourse about "echo chambers" and "filter bubbles" on the internet [10, 11], the available evidence on the precise scope and magnitude of those borders remains ambivalent [12, 13]. It is still unclear how search results differ across geographic, linguistic, cultural, political, and other borders [14].
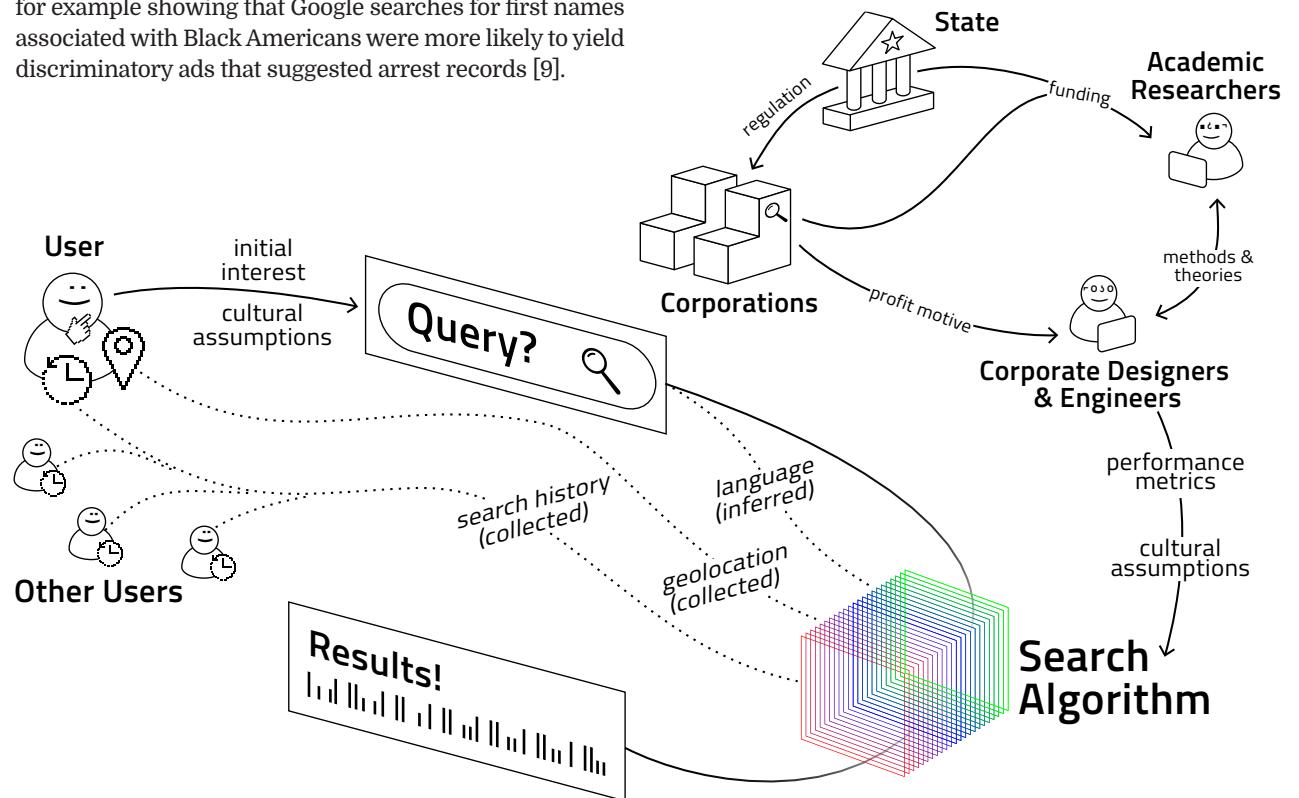
Fig. 2. Search results are products not only of the users' own interests but also of complex struggles among the state, corporations, and academia. Search algorithms encode the cultural assumptions and performance metrics of the designers and engineers, which in turn are shaped by corporate profit motives, state regulations, and academic methods and theories. For a particular query, the search algorithm may tailor the results according to the user's search history, geolocation, and language, as well as other users' data.

Our project, Search Atlas, offers a critical examination of these information borders. Starting in the late sixteenth century, an "atlas" has meant a collection of maps. Despite their aspirations to surveying the entire world and providing definitive accounts of its geopolitical territories, atlases have always been shaped by the assumptions and interests of their makers. Today's digital maps are no different. For disputed territories, Google Maps shows different maps depending on the location of the viewer. For example, if you view Google Maps from India, the region of Kashmir appears to be part of India, shown with a solid border. But if you view Google Maps from anywhere else in the world, a dotted border near Pakistan makes it clear that Kashmir's ownership is disputed [15].
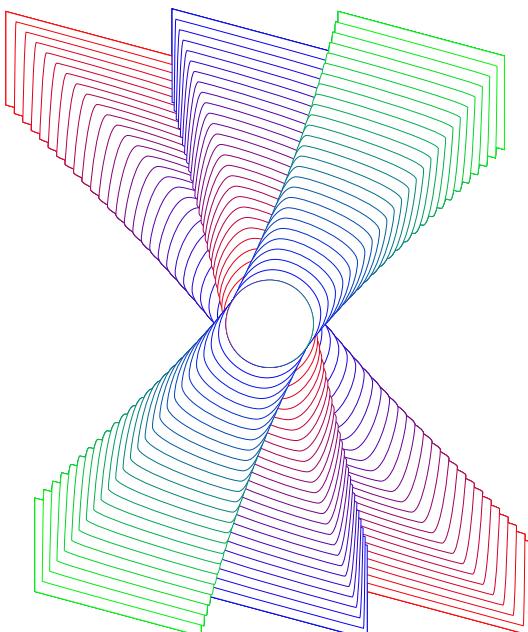
By the mid-nineteenth century, the term "atlas" had spread from geography to the empirical sciences more broadly, ranging from astronomy to botany to anatomy. Atlases became not just collections of maps in the traditional cartographic sense, but "maps" of knowledge in a general sense [16]. If today's search engines are the most extensive and systematic maps of knowledge available, Search Atlas offers a way to compile and compare these maps. But unlike most atlases, Search Atlas does not aim to present an all-encompassing, objective view of the world. Following critical cartographers [17], we appropriate the concept of "atlas" to show that search results are always partial and contested.

In this pictorial, we present a critical intervention in the design of search interfaces, paired with visualizations that expose the partial perspective of the search engine (here, Google). Our project consists of three parts:

1. A tool that enables users to search for any query in any three Google-supported countries (with accompanying languages), returning Google's text and image results for each set of parameters. Users may optionally have their queries machine-translated into each language and the results translated back into their own language. Then, the tool highlights the most distinctive words in each list of results. Example queries are shown in Figs. 3, 4, 5, and 6.

2. A collection of image maps (Figs. 1 and 7) that show the image results for selected queries in almost every Google-supported country. The top images are placed on a tile map in the approximate geographic location of the result's country of origin.

3. A collection of cluster maps (Figs. 8 and 9) that reveal "information regions" and "information borders" in text results for selected queries worldwide. Queries are performed in the Google-determined default language for each country. Then, the results are machine-translated into English, and automatically clustered by text similarity, so countries with similar results are spatially grouped together.

(Due to the cost of running hundreds of queries for each visualization, the image and cluster maps in this pictorial are handmade and not available for user-supplied queries.)

For a demonstration of our tool, see searchatlas.org.

RELATED WORK AND CRITICAL FRAMEWORKS

The concentration of power in technological infrastructures has become a matter of public concern. Such infrastructures, including search engines, seem to play a key role in the spread of false information and hate speech, including the white supremacist and Islamophobic content that has fueled such disastrous incidents as the U.S. Capitol riot and the Rohingya genocide in Myanmar. Scholars at the intersection of science and technology studies and critical race theory have paved the way for understanding the role of technology in these incidents [4, 18, 19]. Our design work is guided by their critiques, as well as by several lines of work that incorporate critical concerns into artistic and technical interventions.
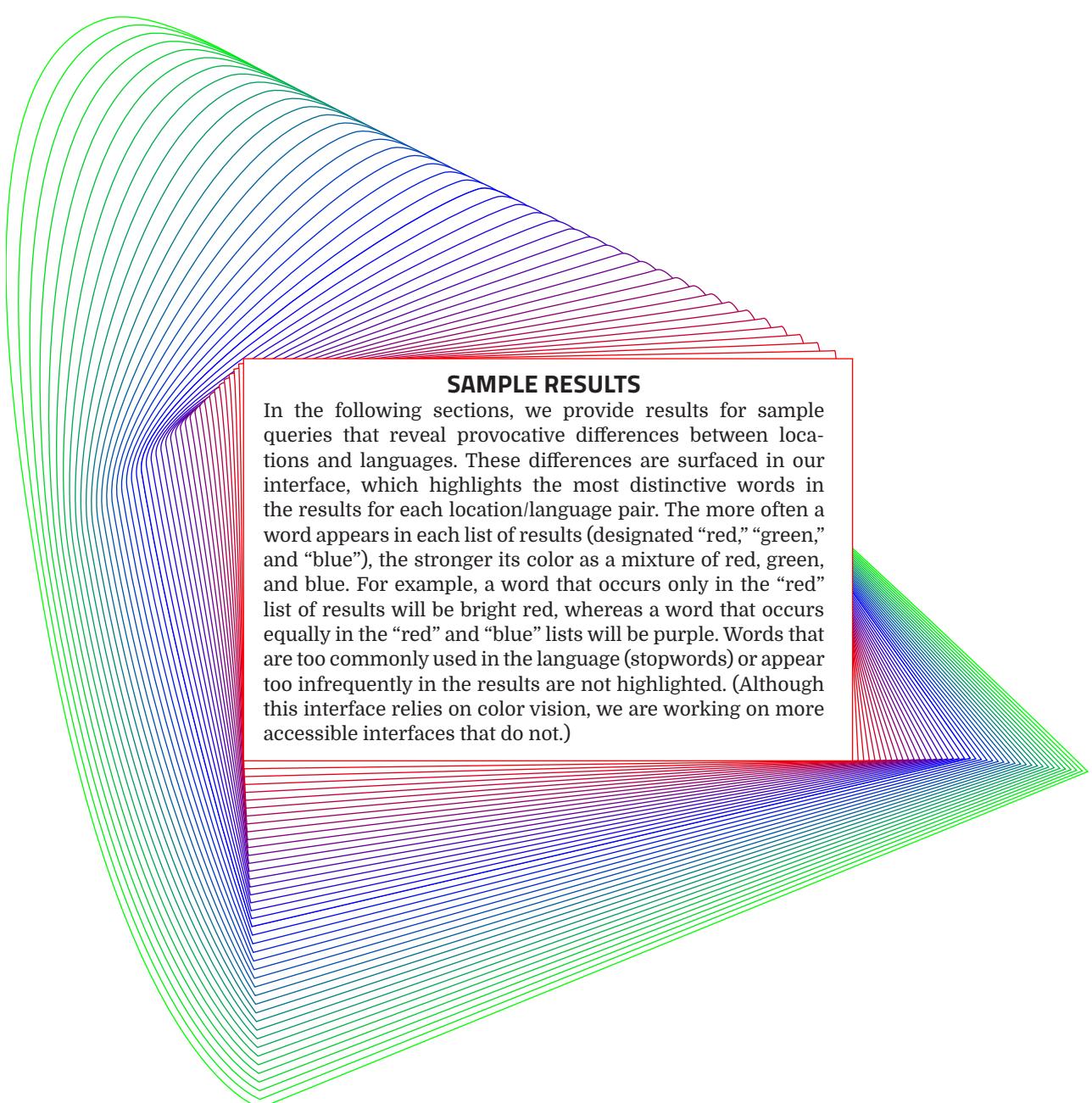
Media art offers precedents for designs that defamiliarize an individual user's experience with a technology in order to make a point about their everyday experience. For example, Taryn Simon and Aaron Swartz's "Image Atlas" (2012) juxtaposes image search results for a single query in multiple search engines in different countries, arranging the results into a list of hundreds of images. Simon and Swartz presented their project as a study of "cultural differences and similarities" across the world and as an investigation of mediation: how "tools like Facebook and Google … are programmed and are programming us" [20]. Recent media interventions have continued to explore ways to transgress the algorithmic filters imposed by online platforms. "PolitEcho" (2017) visualizes the "filter bubble" of one's Facebook friends [21], and Mozilla's "TheirTube" (2020) simulates YouTube video recommendations for users with different cultural profiles, such as "liberal" or "conservative" in the U.S. sense [22]. By juxtaposing divergent views, our interface follows these tactics for making users aware of the layers of mediation behind the information they receive.

However, we present Search Atlas not just as an art piece that facilitates a one-time experience, but as a tool that could be plausibly used in everyday life. Thus, Search Atlas can be understood as a work of critical design [23, 24, 25]. For example, mainstream search interfaces like Google's make the design assumption of a single language

and a single location per user. Our interface questions the value judgments of this "affirmative design" [26] by supporting user personas that are underserved, such as those of migrant and multilingual populations. By encoding values of cultural and linguistic multiplicity, Search Atlas invites users to speculate on the kind of world that would surround it (à la [27]), a world where interfaces that embrace plurality are not the exception but the norm.

Yet our work also seeks to move behind the interface, to probe the technical operations of search algorithms. As the Critical Engineering Manifesto puts it, "The greater the dependence on a technology the greater the need to study and expose its inner workings" [28]. Could internet users be any more dependent on search engines? Putting this manifesto into action, the Critical Engineering Working Group produces custom software that exposes the inner workings of widespread but ill-understood technologies. While, in our setting, it is not possible to "open the black box" of proprietary search engines as one can open a cell phone, we study the workings of search engines scientifically, prodding them with exhaustive combinations of inputs and identifying patterns in their outputs.

Overall, our work aims to "study up": to appropriate the tools of the powerful, which are typically deployed against more vulnerable groups, to instead hold the powerful to account. This tactical move from anthropology is increasingly making its way into computer science as a reaction to the latter's tendency to "study down" [29]. One example of "studying up" computationally is the Dark Inquiry collective's "White Collar Crime Risk Zones" (2017) [30]. Rather than predict "blue-collar" crime as is typical with algorithmic risk assessment, this project uses machine learning to predict where financial, "white-collar" crime is likely to happen and visualizes the results. In what follows, we "study up" by applying standard data analysis techniques employed by search engines (like tf-idf [31]), as well as cutting-edge visualization techniques (like UMAP [32]) to search engine results. Our goal is to open up the search engine to critical interrogation.

SAMPLE RESULTS

In the following sections, we provide results for sample queries that reveal provocative differences between locations and languages. These differences are surfaced in our interface, which highlights the most distinctive words in the results for each location/language pair. The more often a word appears in each list of results (designated "red," "green," and "blue"), the stronger its color as a mixture of red, green, and blue. For example, a word that occurs only in the "red" list of results will be bright red, whereas a word that occurs equally in the "red" and "blue" lists will be purple. Words that are too commonly used in the language (stopwords) or appear too infrequently in the results are not highlighted. (Although this interface relies on color vision, we are working on more accessible interfaces that do not.)

Search Atlas

god | Translate query? ☑

FIG. 3. RESULTS FOR A SEARCH FOR "GOD"

| Google Japan in Japanese ▾ | Google United Arab Emirates in Arabic ▾ | Google United States in English ▾ |

Submit | Translate results? ☑

Results for '神' on Google from Japan in Japanese

God (Shinto) --Wikipedia (ja.wikipedia.org)

A god in Shinto is an object of belief and awe such as natural phenomena. When we say "eight million gods", "eight million" is an analogy of a large number.

A Japanese god directory that Japanese people should know. It changes with the times ... (discoverjapan-web.com)

The two gods who finished the birth of the country will give birth to many gods next, but Izanami will be burned and die when he gives birth to the god of fire, Hino Kagutsuchi. However, the remaining Izanagi continued to give birth to a god, such as Amaterasu and Susanoo ...

What is God for Japanese people | nippon.com (www.nippon.com)

Many Japanese people think of religion as "God" or "Buddha". In particular, "God" has been an object of worship even before the introduction of Buddhism. How was God (Kami) considered in Japan from ancient times to modern times?

What is God [Pixiv Encyclopedia] --pixiv (dic.pixiv.net)

A word that expresses a paranormal existence that transcends humans. God is a being that transcends human beings and is an object of worship that gives blessings and punishments to human beings. Appears in the myths and legends of each country. + One of the net slang. With wonderful technology and sensibility that seems to be human work ...

In Japan, the results emphasize Shinto spirits (kami).

Results for 'الله' on Google from United Arab Emirates in Arabic

God - Wikipedia (ar.wikipedia.org)

Allah (the word of majesty). It is a singular science that has no plural in the Arabic language indicating the "deity" and "creator" in the divine monotheistic religions (and other belief systems). "God" is described as ...

Allah (Islam) - Wikipedia (ar.wikipedia.org)

God in Islam is a science on the self, the duty of existence that is worthy of all the praises, and it is the name of the Sublime Self, the Creator of the universes and existence, and He is the true God of all creatures, and there is no rightly worshiped but He.

God - knowledge (www.marefa.org)

In Islam, God is the one true, supreme God, the All-Powerful, the All-Knowing, the Creator, the Provider, the Al-Kafi, the owner of the universe. Islam places a heavy pillar on elaborating a perception of God with unique precision with regard to ...

What is the answer of the Messenger 🌸 to the question, Where was God Almighty before he was created ... (www.youtube.com)

What is the answer of the Messenger 🌸 to the question, Where was God Almighty before the creation of creatures? | Dr.. Tariq Al-Suwaidan, this video from the first episode of the Stories of the Prophets series, link of the episode ...

In the United Arab Emirates, they point exclusively to Islamic sources.

Results for 'god' on Google from United States in English

God - Wikipedia (en.wikipedia.org)

In monotheistic thought, God is conceived of as the supreme being, creator, and principal object of faith. God is usually conceived of as being omnipotent, omniscient, omnipresent and omnibenevolent as well as having an eternal and necessary existence.

God in Christianity - Wikipedia (en.wikipedia.org)

God in Christianity is the eternal being who created and preserves all things. Christians believe God to be both transcendent and immanent (involved in the ...

Who Is God? Explore the God of The Bible | BibleProject™ (bibleproject.com)

Explore God and the question "who is God" with BibleProject™. Learn and understand the complex identity of ...

God | Definition of God by Merriam-Webster (www.merriam-webster.com)

1 God : the supreme or ultimate reality: such as. a : the Being perfect in power, wisdom, and goodness who is worshipped (as in Judaism, Christianity, Islam, and Hinduism) as creator and ruler of the universe Throughout the patristic and medieval periods, Christian theologians taught that God created the universe ...—

In the United States, they refer exclusively to a monotheistic Christian god.

Search Atlas

Crimean annexation | Translate query? ☑

FIG. 4. RESULTS FOR A SEARCH
FOR "CRIMEAN ANNEXATION"

| Google Russian Federation in Russian ⌄ | Google Ukraine in Ukrainian ⌄ | Google Netherlands in Dutch ⌄ |

Submit | Translate results? ☑

Results for 'Крымская аннексия' on Google from Russian Federation in Russian

Accession of Crimea to the Russian Federation ... (ru.wikipedia.org)

Ukraine and the countries of Europe and the West regard the events of February - March 2014 as the annexation of Crimea by the Russian Federation, while with ...

The problem of Crimea belonging - Wikipedia (ru.wikipedia.org)

The problem of Crimea's ownership - the disagreements between Russia and Ukraine ... and the subsequent illegal annexation of [Crimea] by the Russian Federation "cannot be the basis for any change in the status of the ARC and Sevastopol.

Annexation of Crimea | News from Germany about Ukraine | DW ... (www.dw.com)

Annexation of Crimea. An illegal referendum on the status of the Ukrainian peninsula was held on March 16, 2014 at gunpoint ...

The annexation of Crimea: German experts on Crimean rhetoric ... (www.dw.com)

The President of Ukraine stands for an international platform for resolving the issue of the annexation of Crimea by Russia. German experts do not ...

Results for 'Кримська анексія' on Google from Ukraine in Ukrainian

Annexation of Crimea (2014) - Wikipedia (uk.wikipedia.org)

This article describes both the illegal annexation itself and the process of Russia's occupation of Crimea. Temporary occupation of the Autonomous Republic of Crimea and Sevastopol ...

Crimean crisis - Wikipedia (uk.wikipedia.org)

Annexation of the Crimean Khanate · Crimean crisis (1992–1994) · Conflict over the island of Tuzla · Occupation of Crimea by Russia (2014) · Russian intervention in ...

Occupation and annexation of Crimea by Russia - Radio Svoboda (www.radiosvoboda.org)

In February 2014, armed people in uniform without identification appeared in the Crimea, who seized the building of the Verkhovna Rada of the Crimea, ...

Annexation of Crimea News - current events about ... (www.dw.com)

This page contains materials published on the Deutsche Welle website that match the search query "Annexation of Crimea".

Results for 'Inlijving op de Krim' on Google from Netherlands in Dutch

Illegal annexation of Crimea and Sevastopol: EU extends ... (www.consilium.europa.eu)

The Council decides the sanctions imposed following the illegal annexation of Crimea and Sevastopol by the Russian Federation until ...

Illegal annexation of Crimea and Sevastopol: EU extends sanctions by ... (www.consilium.europa.eu)

The Council extended the restrictive measures following the illegal annexation of Crimea and Sevastopol by Russia until ...

Sanctions regulation incorporation Crimea and ... --wet.nl - Regulation (wetten.overheid.nl)

MinBuZa.2014.59695, containing restrictive measures in connection with the illegal annexation of Crimea and Sevastopol (Sanctions regime ...

EU extends sanctions against Russia for annexation of Crimea ... (www.europa-nu.nl)

On 20 June 2019, the Council adopted restrictive measures following the illegal annexation of Crimea and Sevastopol by Russia ...

In Russia, the results frame the issue in terms of whether Crimea belongs to the "Russian Federation."

In Ukraine, they frame the issue as an "occupation."

In the Netherlands, they focus on the European Union's sanctions on Russia.

Search Atlas

god [ ] Translate query? ☑

FIG. 5. IMAGE RESULTS
FOR A SEARCH FOR "GOD"

| Google Bulgaria in Bulgarian ⌄ | Google Azerbaijan in Azerbaijani ⌄ | Google Mongolia in Mongolian ⌄ |
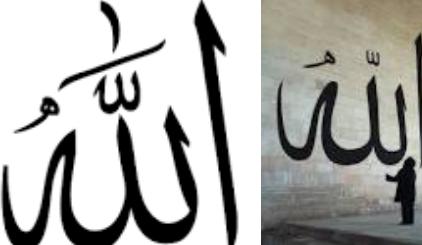
Submit   Translate results? ☑

Results for 'Бог' on Google from Bulgaria in Bulgarian

Results for 'Allah' on Google from Azerbaijan in Azerbaijani

Results for 'Бурхан' on Google from Mongolia in Mongolian
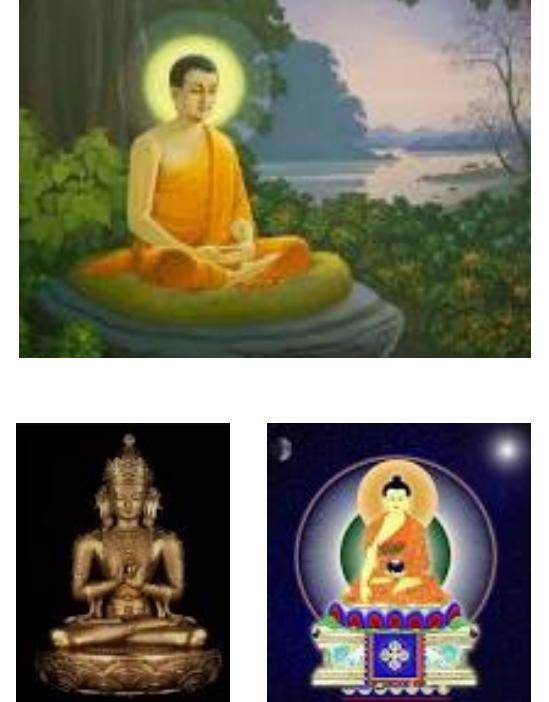
In Bulgaria, the results depict a traditional Christian god.

In Azerbaijan, they are calligraphic images of the word "Allah" in Arabic.

In Mongolia, they are Buddhist paintings.

1976

Search Atlas

Tiananmen Square | Translate query? ☑

FIG. 6. IMAGE RESULTS FOR A SEARCH FOR "TIANANMEN SQUARE"

Google United Kingdom in English | ⌄

Google Singapore in English | ⌄

Google China in Chinese (Simplified) | ⌄

Submit | Translate results? ☑

Results for 'Tiananmen Square' on Google from United Kingdom in English

Results for 'Tiananmen Square' on Google from Singapore in English

Results for '天安门广场' on Google from China in Chinese (Simplified)

In the United Kingdom and in Singapore, the results surface photographs of tanks and soldiers in the 1989 protests, which were widely circulated in the international press.

In mainland China, they present touristic and promotional images. (Google is blocked by the Chinese government, but it is possible to circumvent the block and use Google with China as the country setting.)

1977

IMAGE SEARCH AROUND THE WORLD

Examining the results for one query in a few carefully chosen locations gave us an incisive picture of just how different the results can be on the ground. The next step is to widen the field of view. Here, we provide visualizations of search results for sample queries worldwide.

To create these visualizations, we machine-translate each query into the default language for each country using Google Translate. To determine the default languages, we scrape the `google.com` homepage with each possible country parameter (`gl`) and detect the language parameter (`hl`) set as default for each.
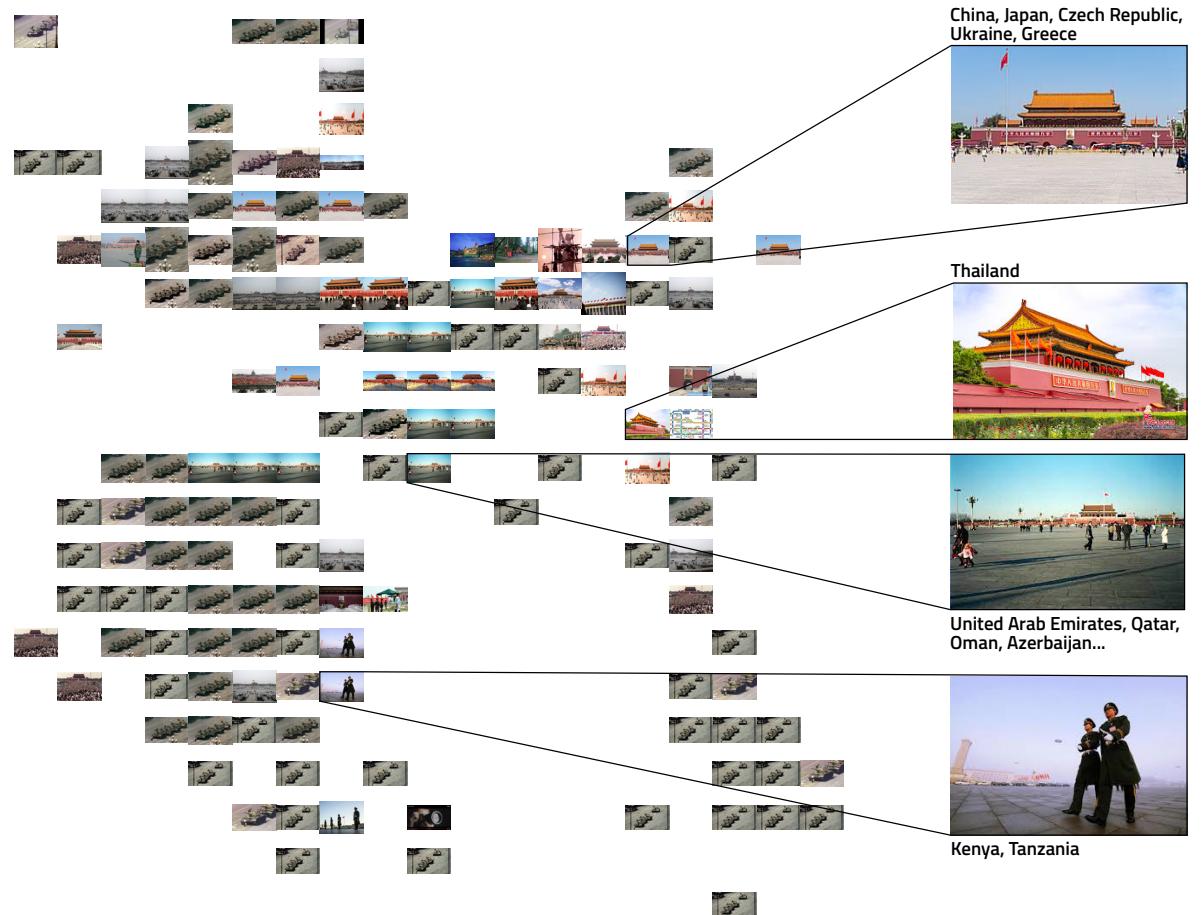
It is important to note that Google tends to select state-sanctioned and colonial languages as the default for a country. The default language for Mali is French, which is the state's official language, even though Bambara is much more widely spoken. English is the default language for Pakistan even though Urdu is also official and others, such as Punjabi, are more widely spoken.

We provide two image maps, one in Fig. 1 (in the title page, for "god") and the other in Fig. 7 (below, for "Tiananmen Square").


FIG. 7. A WORLD MAP OF TOP IMAGE RESULTS FOR SEARCHES FOR "TIANANMEN SQUARE"

This world map shows top image results for searches for "Tiananmen Square" made in most Google-supported countries, overlaid on the approximate geographic location of each country. Most countries show photographs of the 1989 protests, except for China and a few surrounding countries, which show touristic and promotional images.

MAPPING INFORMATION REGIONS

Looking at the image maps, we readily notice ways to group countries. For example, when it comes to "god," countries as far apart as Bhutan and Gibraltar lie in the same "information region": searchers in both countries would find similar images of a Western god looking at a kneeling Jesus. On the other hand, countries as close as Egypt and Sudan lie in different information regions: searchers in Sudan would find the same Christian image, whereas searchers in Egypt would find a calligraphic representation of Allah. Regardless of the underlying causes, there is some kind of "information border" between these regions.
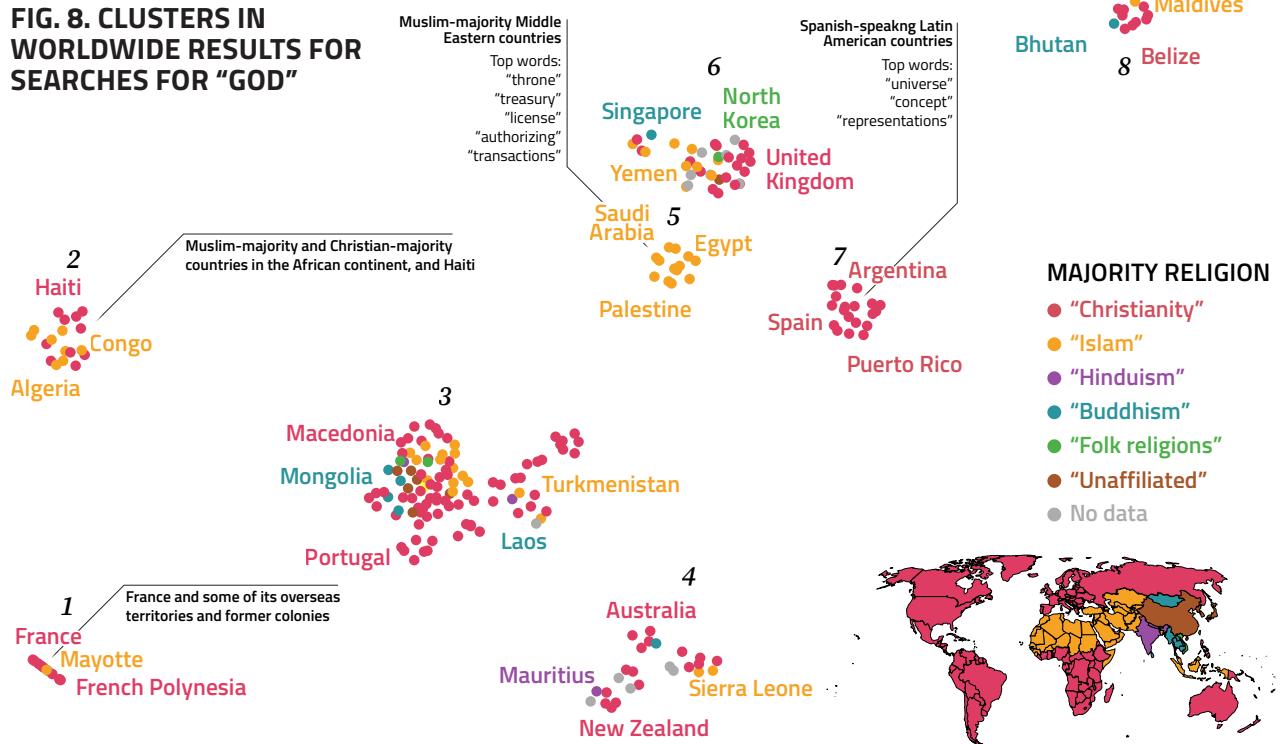
To identify such information regions and borders more precisely, we choose to analyze text results, not images, since there are better-defined ways of measuring text similarity. We analyze the text using automatic methods that reveal which countries have similar and dissimilar results, enabling us to remap the world by the similarity of the information that a searcher sees.

Specifically, we again make searches worldwide using the Google-determined default language for the country. Then, given the text results in that language, we machine-translate the results back into English, again using Google Translate. Each country's English results can be understood as an approximately hundred-dimensional vector of its most distinctive words, found via the tf-idf algorithm [31]. (For example, Japan's top words in its results for "god" are "japanese," "shinto," "kami" [spirits], and "awe.") The similarity between two countries is quantified as the cosine similarity between their vectors. Finally, we use an algorithm called UMAP, which is state-of-the-art for dimensionality reduction [32], to arrange the countries in a two-dimensional space and automatically cluster them according to how similar their search results are.

Note that each map we give is just one of many possible maps, since UMAP is a nondeterministic algorithm whose outputs depend on parameters related to the desired amount of global or local structure to visualize in the data. We choose parameters that lean toward preserving more global structure. Most importantly, the clusters in our analysis appear to persist throughout many runs of the algorithm.

FIG. 8. CLUSTERS IN WORLDWIDE RESULTS FOR SEARCHES FOR "GOD"

Visualization of worldwide results for searches for "god" using UMAP. Each country is colored according to the religion with the largest share of followers in it, as indicated by the map and legend on the right. The religion data were compiled in 2010 by the Pew Research Center [33]. Note that the Pew data use questionable categories, such as "folk religions," and reproduce the nineteenth-century European myth of "world religions" [34].
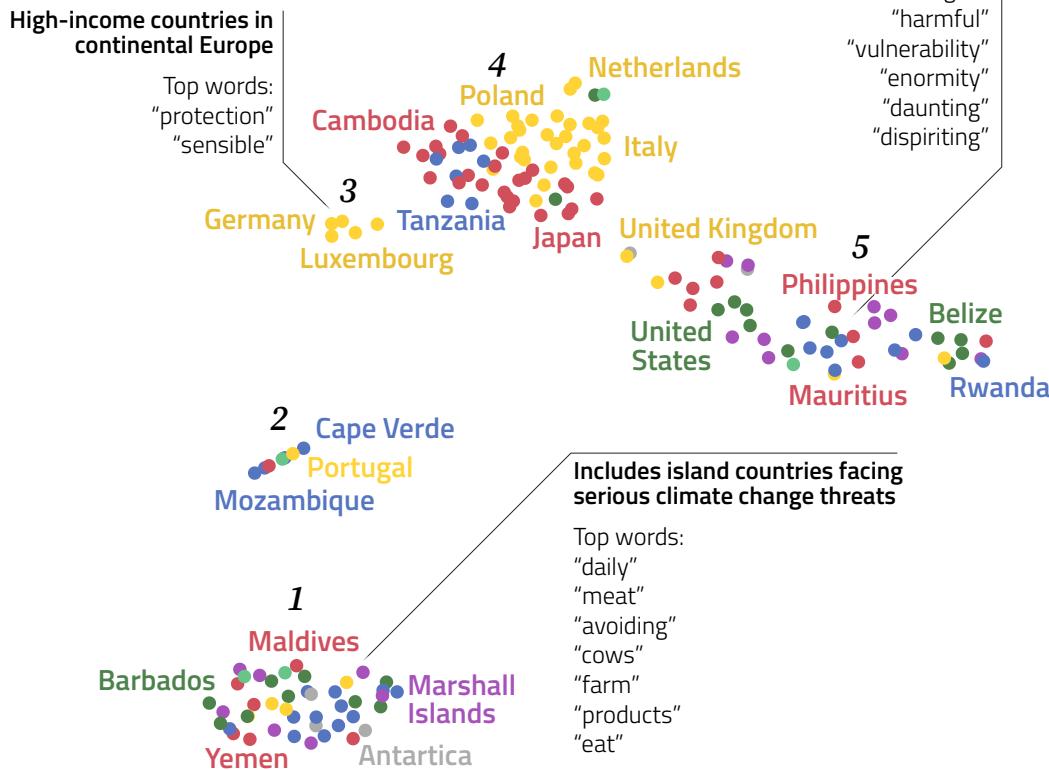
Several clear clusters emerge, which seem to be formed by a combination of common geographic location, language, and religion. Cluster 7 consists of Spanish-speaking Latin American countries, which also tend to be predominantly Christian. Top words include "universe," "concept," "representations." Cluster 5 consists of Muslim-majority Middle Eastern countries. Top words include "throne," "treasury," "license," "authorizing," "transactions."

Francophone countries split into two separate clusters. Cluster 2 consists almost entirely of countries in the African continent, whether Muslim-majority or Christian-majority. Cluster 1 comprises France and its overseas territories and former colonies outside of Africa, ranging from the Caribbean Sea to the Atlantic Ocean to Polynesia. Yet there are exceptions: Haiti is in cluster 2 despite being a Caribbean country.

Cluster 3 is intriguing because it is not easily legible through any of the lenses of geography, language, or religion. It includes countries as diverse as Macedonia, Turkmenistan, and Laos. While this cluster did not have a consistent set of shared top words, the UMAP algorithm judged them to be more similar to each other than to countries in other clusters. This automatically discovered cluster may comprise a new information region.
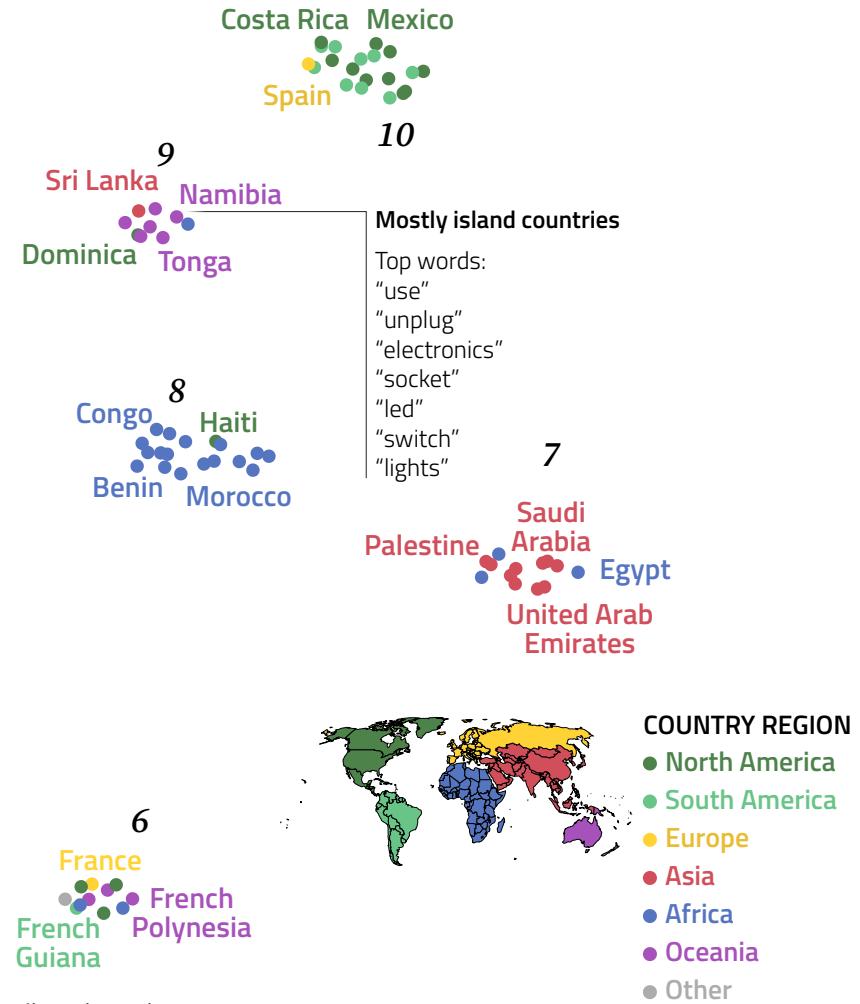
FIG. 9. CLUSTERS IN WORLDWIDE RESULTS FOR SEARCHES FOR "HOW TO COMBAT CLIMATE CHANGE"

High-income countries in continental Europe

Top words:
"protection"
"sensible"

Top words:
"vital"
"signs"
"harmful"
"vulnerability"
"enormity"
"daunting"
"dispiriting"

Mostly island countries

Top words:
"use"
"unplug"
"electronics"
"socket"
"led"
"switch"
"lights"

Includes island countries facing serious climate change threats

Top words:
"daily"
"meat"
"avoiding"
"cows"
"farm"
"products"
"eat"

COUNTRY REGION
- North America
- South America
- Europe
- Asia
- Africa
- Oceania
- Other

Visualization of worldwide results for searches for "how to combat climate change" using UMAP. Each country is colored according to its continent, as indicated by the map and legend on the right [35]. Note that the clusters are more numerous and different than those in the previous example.

The information borders of climate change seem to be defined largely along island versus continental lines. In cluster 3, comprising high-income countries in continental Europe, such as Germany, Liechtenstein, and Luxembourg, the top words suggested preemptive measures on "climate protection" ("protection," "sensible"). Yet, the top words in a few island countries that form part of cluster 5, ranging from Mauritius in the Indian Ocean to Trinidad and Tobago in the Caribbean Sea, suggested much greater immediate threats ("vital," "signs," "harmful," "vulnerability," "enormity," "daunting," "dispiriting").

Some countries' results tended to focus on governmental and institutional policy, while others emphasized individualistic action. The results of cluster 3 included sites of government organizations such as the German federal cabinet and the European Environment Agency. By contrast, the top words in the Netherlands, Aruba, and Suriname focused on consumer choices ("buy," "consume"). In cluster 9, comprising mostly island countries, from Tonga to Sri Lanka, the top words suggested other household practices ("use," "unplug," "electronics," "wall," "socket," "led," "switch," "lights"). In cluster 1, which includes some of the island countries most threatened by climate change,

including Maldives and the Marshall Islands, the top words focused specifically on food choices ("daily," "meat," "avoiding," "cows," "farm," "top," "products," "eat").

These findings are consistent with ethnographic studies of climate change discourses in the Marshall Islands, which have reported that "Despite awareness of their tiny carbon footprint, grassroots Marshall Islanders (if not their government) have strongly favored a response of guilt and atonement rather than outrage and protest" [36].

PRACTICAL LIMITATIONS

The implementation of our tool faced many obstacles. Google intentionally lacks an API for web search results, and deploys various tactics to block scrapers. Ironically, Google is a scraper itself, and profits massively off its scraping of sites. In fact, many sites block all crawlers except Google, due to bandwidth and other capacity constraints that could cause the sites to crash. This behavior creates a barrier to competition by alternative search engines. In the United States and in the European Union, antitrust regulators have cited this barrier among Google's multiple measures to protect its monopoly power. Such measures also include contractual arrangements to make Google the default search engine in major web browsers and operating systems [37].

For researchers, the only legally safe and technically feasible way to obtain data of search results is to use a third-party scraper. We obtained data from an API operated by a third-party firm, which provides defense against Google's legal threats and technical methods to block scrapers. Even then, our data collection has been costly, demanded a substantial amount of code, and faced several other obstacles. For instance, we could not obtain reliable search results from Myanmar. Although we could not determine the precise reason, one possible explanation was an ongoing military coup (at the time of writing) that involved a nationwide internet shutdown. We also could not include Botswana and Seychelles because Google Translate does not support their default languages, Setswana and Seychellois Creole respectively.

THE CHALLENGE OF INTERPRETATION

The internet is full of information borders. Sometimes those borders follow foreseeable lines, for example according to geography and language. But there are frequent exceptions. Many information borders are unpredictable and harder to explain. Geographically close countries with the same language may have significantly different results, and distant countries with unrelated languages may have unexpectedly similar results. By automatically comparing and clustering Google results across countries, we discover that information borders are often surprising and unique. They can vary widely depending on the query, and can transcend geographic, linguistic, and other traditional borders.

How to interpret those new information regions and borders? The underlying infrastructure offers few answers. To return results, Google Search relies heavily on novel deep learning systems whose decisions are notoriously difficult to interpret, even by Google's own researchers [38, 39]. Moreover, our results for information borders and regions are limited by the difficulty of comparing results across geopolitical borders. Our results may be influenced by confounding factors related to language, which are hard to account for in this kind of global investigation, or by mediating factors such as Google's machine translation and the UMAP algorithm.

To get feedback on how people receive our results, we shared early prototypes of Search Atlas in participatory workshops with computer scientists, artists, and designers. In these sessions, we noticed two common temptations in interpreting the results of the tool. One is to interpret the results as straightforward reflections of cultural differences among users in different countries. Another is to interpret the results as unambiguous outcomes of political bias or manipulation by algorithm designers. We encourage our readers to resist both of these temptations. Search engines are not entirely neutral conduits that respond to users' interests with objectively "relevant" results, nor are they reducible to purely subjective editors with unlimited power to pick and choose [40]. Rather, our results underscore how search engines are products both of cultural patterns and of algorithm design choices.

Search engines respond both to users' immediate interests and to corporations' financial imperatives. The design of Google's search engine is inseparable from the priorities of its advertising business [41]. Search engines also respond to political pressures and legal regulations. China-based search engine Baidu favors results that align with the views of Chinese government authorities [42], and Google removes results to comply with European data protection laws [43].
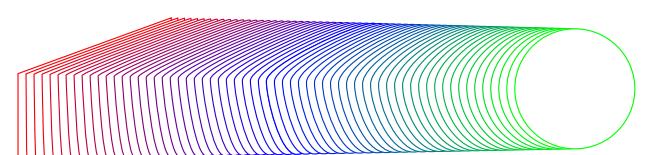
Moreover, search results are products not only of algorithm design but also of human judgment and curatorial labor. Google employs subcontracted workers, such as "raters" who judge the perceived quality of search results and "content moderators" who judge whether results seem inappropriate or illicit [44, 45, 46]. The production of search results also involves other actors with competing interests and goals, such as "search engine optimization" (SEO) consultants who deploy various tactics to help their clients compete for attention [47].

Finally, designers in different places also have different cultural assumptions, concerns, and practices, all of which shape the design of their search engines. According to the co-founders of Yandex, now Russia's most popular search engine, their initial demonstration product was an algorithm for searching a Russian version of the Bible, later adapted for searching the web [48].

In revealing these information borders, we encourage our readers to consider the invisible processes and complex struggles behind the production of search results. Our project invites users to experience the internet from divergent positions and to reflect on how their online lives are conditioned by technological infrastructures and geopolitical regimes.

We hope that our work encourages researchers and activists to experiment with critical interventions in the design of search interfaces, and to study how the everyday use of more pluralistic interfaces could affect people's online lives. Ultimately, we hope that such projects can expose and contest the world's information borders, and perhaps even help to collectively reshape them in more democratic ways.

REFERENCES

[1] How Google Search Works. *Google*. Retrieved from https://www.google.com/search/howsearchworks/

[2] Donna Haraway. 1988. Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective. *Feminist Studies* 14, 3 (1988), 575–599. DOI: https://doi.org/10.2307/3178066

[3] Lucas D. Introna and Helen Nissenbaum. 2000. Shaping the Web: Why the Politics of Search Engines Matters. *The Information Society* 16, 3 (July 2000), 169–185. DOI: https://doi.org/10.1080/01972240050133634

[4] Safiya Umoja Noble. 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York University Press, New York, NY.

[5] Eszter Hargittai. 2007. The Social, Political, Economic, and Cultural Dimensions of Search Engines: An Introduction. *Journal of Computer-Mediated Communication* 12, 3 (April 2007), 769–777. DOI: https://doi.org/10.1111/j.1083-6101.2007.00349.x

[6] Laura A. Granka. 2010. The Politics of Search: A Decade Retrospective. *The Information Society* 26, 5 (September 2010), 364–374. DOI: https://doi.org/10.1080/01972243.2010.511560

[7] Frank Pasquale. 2015. *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard University Press, Cambridge, MA.

[8] Liwen Vaughan and Yanjun Zhang. 2007. Equal Representation by Search Engines? A Comparison of Websites Across Countries and Domains. *Journal of Computer-Mediated Communication* 12, 3 (April 2007), 888–909. DOI: https://doi.org/10.1111/j.1083-6101.2007.00355.x

[9] Latanya Sweeney. 2013. Discrimination in Online Ad Delivery. *Communications of the ACM* 56, 5 (May 2013), 44–54. DOI: https://doi.org/10.1145/2447976.2447990

[10] Cass R. Sunstein. 2009. *Republic.com 2.0*. Princeton University Press, Princeton, NJ.

[11] Eli Pariser. 2011. *The Filter Bubble: What the Internet Is Hiding From You*. Penguin Press, New York, NY.

[12] Aniko Hannak, Piotr Sapiezynski, Arash Molavi Kakhki, Balachander Krishnamurthy, David Lazer, Alan Mislove, and Christo Wilson. 2013. Measuring Personalization of Web Search. *In Proceedings of the 22nd International Conference on World Wide Web (WWW '13)*, Association for Computing Machinery, New York, NY, 527–538. DOI: https://doi.org/10.1145/2488388.2488435

[13] Seth Flaxman, Sharad Goel, and Justin M. Rao. 2016. Filter Bubbles, Echo Chambers, and Online News Consumption. *Public Opinion Quarterly* 80, S1 (2016), 298–320. DOI: https://doi.org/10.1093/poq/nfw006

[14] René König and Miriam Rasch (Eds.). 2014. *Society of the Query Reader: Reflections on Web Search*. Institute of Network Cultures, Amsterdam, Netherlands.

[15] Greg Bensinger. 2020. Google Redraws the Borders on Maps Depending on Who's Looking. *The Washington Post*. Retrieved from https://www.washingtonpost.com/technology/2020/02/14/google-maps-political-borders/

[16] Lorraine Daston and Peter Galison. 2007. *Objectivity*. Zone Books, New York, NY.

[17] kollektiv orangotango+ (Ed.). 2018. *This Is Not an Atlas: A Global Collection of Counter-Cartographies*. transcript Verlag, Bielefeld, Germany.

[18] Simone Browne. 2015. *Dark Matters: On the Surveillance of Blackness*. Duke University Press, Durham, NC.

[19] Ruha Benjamin. 2019. *Race After Technology: Abolitionist Tools for the New Jim Code*. Polity, Medford, MA.

[20] Taryn Simon and Aaron Swartz. 2012. Image Atlas. *Net Art Anthology*. Retrieved from https://anthology.rhizome.org/image-atlas

[21] Sunny He, Zachary Liu, Vivian Mo, and Jonathan Zong. PolitEcho. Retrieved from https://politecho.org/

[22] Tomo Kihara. TheirTube. Retrieved from https://their.tube

[23] Jeffrey Bardzell and Shaowen Bardzell. 2013. What Is "Critical" About Critical Design? In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '13), Association for Computing Machinery, New York, NY, 3297–3306. DOI: https://doi.org/10.1145/2470654.2466451

[24] James Pierce, Phoebe Sengers, Tad Hirsch, Tom Jenkins, William Gaver, and Carl DiSalvo. 2015. Expanding and Refining Design and Criticality in HCI. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (CHI '15), Association for Computing Machinery, New York, NY, 2083–2092. DOI: https://doi.org/10.1145/2702123.2702438

[25] Netta Iivari and Kari Kuutti. 2017. Critical Design Research and Information Technology: Searching for Empowering Design. In *Proceedings of the 2017 Conference on Designing Interactive Systems* (DIS '17), Association for Computing Machinery, New York, NY, 983–993. DOI: https://doi.org/10.1145/3064663.3064747

[26] Anthony Dunne and Fiona Raby. 2001. *Design Noir: The Secret Life of Electronic Objects*. Birkhäuser, Basel, Switzerland.

[27] Richmond Y. Wong, Vera Khovanskaya, Sarah E. Fox, Nick Merrill, and Phoebe Sengers. 2020. Infrastructural Speculations: Tactics for Designing and Interrogating Lifeworlds. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (CHI '20), Association for Computing Machinery, New York, NY, 1–15. DOI: https://doi.org/10.1145/3313831.3376515

[28] Julian Oliver, Gordan Savičić, and Danja Vasiliev. 2011. The Critical Engineering Manifesto. *ZKM*. Retrieved from https://zkm.de/en/the-critical-engineering-manifesto

[29] Chelsea Barabas, Colin Doyle, JB Rubinovitz, and Karthik Dinakar. 2020. Studying Up: Reorienting the Study of Algorithmic Fairness Around Issues of Power. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (FAT* '20), Association for Computing Machinery, New York, NY, 167–176. DOI: https://doi.org/10.1145/3351095.3372859

[30] Brian Clifton, Sam Lavigne, and Francis Tseng. 2017. White Collar Crime Risk Zones. *The New Inquiry*. Retrieved from https://whitecollar.thenewinquiry.com

[31] Karen Spärck Jones. 2004. A Statistical Interpretation of Term Specificity and Its Application in Retrieval. *Journal of Documentation* 60, 5 (October 2004), 493–502. DOI: https://doi.org/10.1108/00220410410560573

[32] Leland McInnes, John Healy, and James Melville. 2020. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *arXiv:1802.03426* [cs, stat] (September 2020). Retrieved from http://arxiv.org/abs/1802.03426

[33] Main Religion of the Country, 2010. *Our World in Data*. Retrieved from https://ourworldindata.org/grapher/main-religion-of-the-country-in

[34] Tomoko Masuzawa. 2005. *The Invention of World Religions, or, How European Universalism Was Preserved in the Language of Pluralism.* University of Chicago Press, Chicago, IL.

[35] Continents according to Our World in Data. *Our World in Data*. Retrieved from https://ourworldindata.org/grapher/continents-according-to-our-world-in-data

[36] Peter Rudiak-Gould. 2014. Climate Change and Accusation: Global Warming and Local Blame in a Small Island State. *Current Anthropology* 55, 4 (August 2014), 365–386. DOI: https://doi.org/10.1086/676969

[37] Cecilia Kang, David McCabe, and Daisuke Wakabayashi. 2020. U.S. Accuses Google of Illegally Protecting Monopoly. *The New York Times*. Retrieved from https://www.nytimes.com/2020/10/20/technology/google-antitrust.html

[38] Pandu Nayak. 2019. Understanding Searches Better Than Ever Before. *Google*. Retrieved from https://blog.google/products/search/search-language-understanding-bert/

[39] Tolga Bolukbasi, Adam Pearce, Ann Yuan, Andy Coenen, Emily Reif, Fernanda Viégas, and Martin Wattenberg. 2021. An Interpretability Illusion for BERT. *arXiv:2104.07143* [cs] (April 2021). Retrieved from http://arxiv.org/abs/2104.07143

[40] James Grimmelmann. 2014. Speech Engines. *Minnesota Law Review* (January 2014). Retrieved from https://scholarship.law.umn.edu/mlr/299

[41] Bernhard Rieder and Guillaume Sire. 2014. Conflicts of Interest and Incentives to Bias: A Microeconomic Critique of Google's Tangled Position on the Web. *New Media & Society* 16, 2 (March 2014), 195–211. DOI: https://doi.org/10.1177/1461444813481195

[42] Min Jiang. 2014. The Business and Politics of Search Engines: A Comparative Study of Baidu and Google's Search Results of Internet Events in China. *New Media & Society* 16, 2 (March 2014), 212–233. DOI: https://doi.org/10.1177/1461444813481196

[43] Astrid Mager. 2017. Search Engine Imaginary: Visions and Values in the Co-Production of Search Technology and Europe. *Social Studies of Science* 47, 2 (April 2017), 240–262. DOI: https://doi.org/10.1177/0306312716671433

[44] Lilly Irani. 2015. Justice for "Data Janitors." *Public Books*. Retrieved from https://www.publicbooks.org/justice-for-data-janitors/

[45] Tarleton Gillespie. 2018. *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media.* Yale University Press, New Haven, CT.

[46] Sarah T. Roberts. 2019. *Behind the Screen: Content Moderation in the Shadows of Social Media.* Yale University Press, New Haven, CT.

[47] Malte Ziewitz. 2019. Rethinking Gaming: The Ethical Work of Optimization in Web Search Engines. *Social Studies of Science* 49, 5 (October 2019), 707–731. DOI: https://doi.org/10.1177/0306312719865607

[48] Marina Fedorova. 2019. From Lurker to Ninja: Creating an IT Community at Yandex. In *From Russia with Code: Programming Migrations in Post-Soviet Times*. Duke University Press, Durham, NC, 59–86. DOI: https://doi.org/10.1215/9781478003342-003