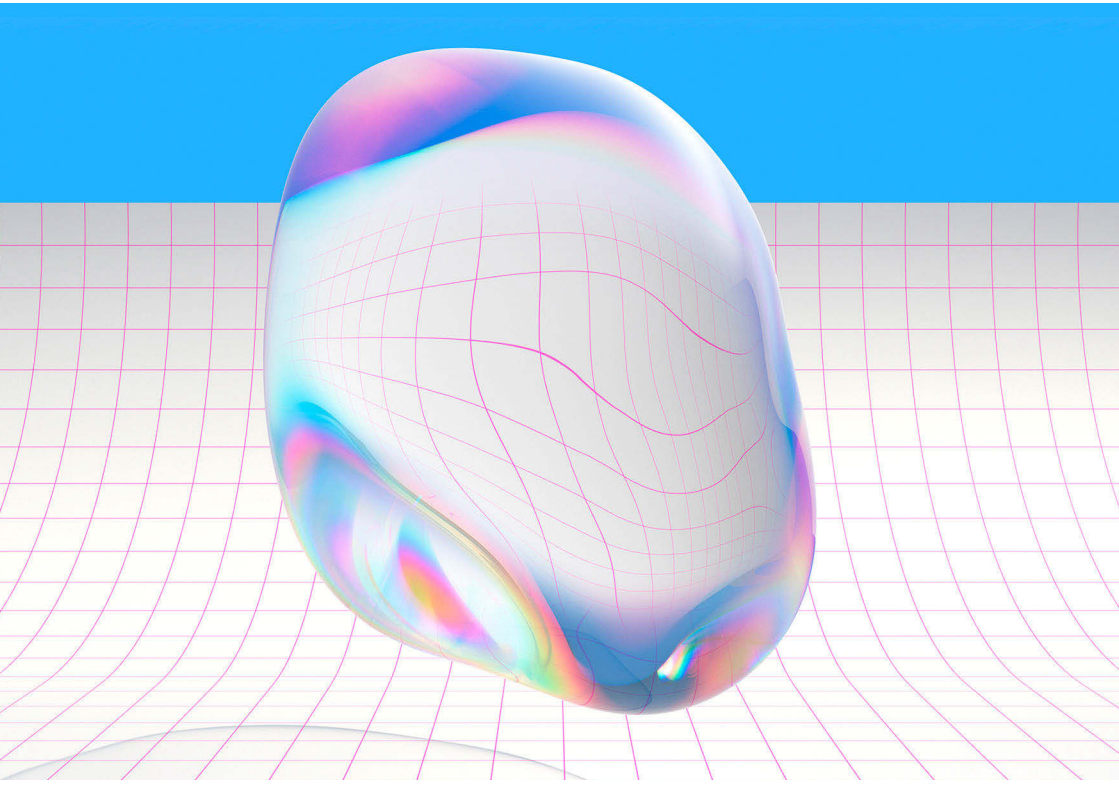
How open is it?

Understand the issue: Openness



The internet is transformative because it is open: Everyone can participate and innovate. But openness is not guaranteed – it's always under attack.

https://doi.org/10.14361/9783839449462-004 https://www.inlibra.com/de/agb - Open Access -

Openness is a foundational pillar of the internet. Today's digital world exists because people don't need permission to create for and on the Web.

Yet in 2019, the internet's openness is as radical – and as threatened – as ever.

Governments worldwide continue to restrict internet access in a multitude of ways, ranging from outright censorship to requiring payment of taxes to use social media, to shutting down or slowing down the internet to silence dissent. Powerful lobbyists are winning fights for more restrictive copyright regimes, and big tech platforms lock us in to proprietary systems

At the same time, the open Web is resilient.

Volunteers of the Wikidata community of Wikimedia have created a data structure that enables content to be read and edited by both humans and machines. Advocates of open data are pushing for more transparency to understand how companies create digital profiles of us and what they do with the data.

But a tension between openness and inclusion persists. Despite many measures taken, hate speech and harassment on online platforms remains an urgent and serious problem.

In Germany, one year after implementation, a law to reduce hate speech online, was neither particularly effective at solving what it set out to do, nor as restrictive as many feared.

Yet the lack of strong evidence isn't stopping similar regulations from being introduced elsewhere. The European Union is currently debating new rules that would require companies of all sizes to take down 'terrorist content' within one hour, or face stiff penalties.

Opponents warn that the law risks undermining people's fundamental rights and stifling competition by setting a bar only the largest companies can meet.

Heightened discussions about artificial intelligence and automated decision making (AI) are also introducing new angles to this debate.

New user-friendly AI tools have made it easier than ever to create deepfakes: media that depict a person saying or doing something they never did. These sort of developments raise a critical question: how do we mitigate the real harms that misuse of a technology could cause, particularly to vulnerable groups, without sacrificing the benefits of the open internet?

Sometimes, the best approach might be to never release it.

https://doi.org/10.14361/9783839449462-004 https://www.inlibra.com/de/agb - Open Access -

OpenAI recently built a language model so good at automatically generating convincing text that they became concerned about it being misused. To mitigate potential harm, the organization decided to release a limited version of the tool. The choice sparked criticism that it was the "opposite of open," while others praised the decision as a "new bar for ethics."

Grappling with the challenge of safeguarding the open internet, while building an inclusive digital world, remains a pivotal task for companies, technologists, policy makers and citizens alike.

This is especially true as a new dimension emerges, centered around an urgent question: how do we decide what technologies to build and use at all?

Show me my data, and I'll tell you who I am

"Stop manipulating us, and give us real choices," says Katarzyna Szymielewicz, a technology and human rights expert, lawyer and activist who advocates for people to have more control over how their data is processed and used.

Companies are building digital profiles of us, made up of data collected by thousands of trackers in mobile apps or on the web. They gather information about us practically whenever we are connected to the internet. Data brokers sell this data to whoever is willing to pay the price. It changes hands between countless companies without our knowledge.

Data about us is sorted into categories we often can't see and analyzed by algorithms we often don't know about – and then used to make decisions that could impact our lives, for better or worse.

But what if we could take guessing out of the equation, and just *tell* companies who we are? Would they respect our answers?

Katarzyna Szymielewicz is the co-founder and president of Panoptykon Foundation, a digital rights organization in Poland. In January 2019, Panoptykon filed a complaint against Google under new the European General Data Protection Regulation, alleging the company had violated the regulation's requirements to provide users with access to data held about them.

https://doi.org/10.14361/9783839449462-004 https://www.inlibra.com/de/agb - Open Access -

First layer: What you share

The first layer is information we actively feed into social media and mobile applications. We can control this data ourselves if we choose not to share specific information: not to publish certain updates, not to upload photos, avoid sensitive search queries, and so on.

Second layer: What your behaviour tells them

The second layer is our behavioral data and 'metadata' logged by our devices. For example our current location or who we communicate with. It is possible to control this layer of our digital profile to some extent, but it requires real effort and technical expertise.

Third layer: What the machine thinks about you

The third layer is interpretations of the data collected in the first and second layers by algorithms that learn who we are based on behaviors and statistical correlations. It is virtually impossible to control. Full access to data generated by algorithms is often not made available to users.

See interactive visualization on the Internet Health Report 2019 website.

To help a broader audience visualize how little we're currently able to control our digital profiles, Szymielewicz has developed a metaphor of "three layers" of data: providing examples of what is collected about us, what is observed and what is generated by machines.

Q: Are our data profiles inaccurate?

Katarzyna Szymielewicz: Who knows? Without transparency and access to the full profiles that are generated for us by tech companies we cannot really tell. I am sure users themselves would be the best auditors of these datasets because they have real (often economic) incentives not to be judged on the basis of incorrect or incomplete information. But they are not given the chance to do so.

I came up with this layered metaphor to explain the complexity (and dangers) of how online data profiles work after hearing for the hundredth time: 'What's the problem if we choose to share and publish our data ourselves?' The

https://doi.org/10.14361/9783839449462-004 https://www.inlibra.com/de/agb - Open Access -

thing is that we do *not* make these choices ourselves. We are lured into sharing more data than we would accept, observed and qualified by machines in ways we can hardly imagine. Not surprisingly, they detect sensitive characteristics we may prefer to keep private.

Q: Why should we want to see our data?

The only way to regain full control over our profiles, is to convince the companies who do the profiling to change their approach. Instead of hiding our data from us, they should become more transparent. We need to open these opaque systems to the scrutiny of users.

On the other hand – instead of guessing our location, relationships, or hidden desires behind our backs, I think companies could simply start asking us questions, and respecting our answers. I even see this as a real opportunity for marketing companies to build trust and make targeted ads more relevant and fair.

In the European Union, we have a legal framework that facilitates greater openness and access. The General Data Protection Regulation (GDPR) now gives Europeans the right to verify data held by individual companies, including marketing and advertising profiles. Companies can still protect their code and algorithms as business secrets, but in theory they can no longer hide personal data they generate about their users. I say in theory – because in practice companies don't reveal the full picture when confronted with this legal obligation. In particular, they hide behavioural observation data and data generated with proprietary algorithms. This must change, and I am sure it will, once we begin to see the first legal complaints result in fines.

Q: How could we make radical transparency a reality?

Well, no doubt we have to be prepared for a long march. We need to work together as a movement and test different approaches. Some of us will continue to test legal tools and fight opponents in courts or in front of Data Protection Authorities. Others will advocate for (still) better legal safeguards, for example in the upcoming European ePrivacy Regulation. Others will build or crowdfund alternative services or push big tech to test new business models, and so on. I am sure it will be a long run, but as a movement, we are at

https://doi.org/10.14361/9783839449462-004 https://www.inlibra.com/de/agb - Open Access -

least heading in the right direction. The main challenge for us now is to convince or compel commercial actors to come along.

► Further reading

- Networks of Control: A Report on Corporate Surveillance, Digital Tracking, Big Data & Privacy, Wolfie Christl and Sarah Spiekermann, 2016. https://crackedlabs.org/en/networksofcontrol
- Data Ethics – the new competitive advantage, Gry Hasselbalch and Pernille Tranberg, 2016. https://dataethics.eu/wp-content/uploads/DataEthics-UK-original.pdf
- The Age of Surveillance Capitalism by Shoshana Zuboff review – we are the pawns, The Guardian, 2019. https://www.theguardian.com/books/2019/feb/02/age-of-surveillance-capitalism-shoshana-zuboff-review
- Your digital identity has three layers and you can protect only one of them, Quartz, 2019. https://qz.com/1525661/your-digital-identity-has-three-layers-and-you-can-only-protect-one-of-them/

► Further listening

- All Your Data Are Belong to Us, IRL podcast, S.1 E.1, 2017

Internet slowdowns are the new shutdowns

"Imagine you're on a flight. You don't know when you will arrive. You're perpetually stuck in the air until the pilot decides to land." That's how Berhan Taye from Ethiopia describes the strange limbo of an internet shutdown. She leads the #KeepItOn campaign of Access Now, which brings together a coalition of organizations to keep the internet open and accessible.

Around the world, internet shutdowns are on the rise. In 2018, Access Now documented 188 shutdowns around the world. That's more than double what they counted in 2016. Most shutdowns occurred in Africa and Asia, with India being the worst offender.

Official justifications range from cracking down on terrorism, social unrest or false political rumors, to the curbing of cheating during school exams. Other times, authorities simply deny a shutdown or offer no expla-

https://doi.org/10.14361/9783839449462-004 https://www.inlibra.com/de/agb - Open Access -

nation at all. Each case is different, but every time the internet is shut down, peoples' rights are denied. In Cameroon, the government completely blocked Anglophone regions from accessing social media for 230 days. In Chad, citizens have not been able to freely access Whatsapp, Facebook and Twitter for nearly a year. Blocking like this is another way that many governments have been trying to subtly censor access to information without attracting the attention of a full network shutdown.

Recently, Taye says, governments, police and local authorities have become more tactical about how they block people from getting online, moving from internet shutdowns to slowdowns to further obscure who is responsible.

Q: What is it like to experience an internet shutdown?

Berhan Taye: With #KeepItOn we've begun collecting and sharing more personal stories of shutdown experiences because people have a hard time understanding the human impact. Internet shutdowns don't just happen on a random Tuesday. They tend to happen in the context of election violence, protests, or emergencies. That's one reason it can be such a traumatic experience. We've heard from people in the Democratic Republic of Congo, who were unable to verify if family members were alive. We also know of cases in Cameroon where doctors working with the World Health Organization were providing emergency medical advice to patients over WhatsApp. With no internet, there was suddenly no way for them to administer care.

In shutdowns, people's lives are more than just inconvenienced at work, in school, and at home. Their lives can be endangered. In Pakistan, a woman was struck down in traffic by a hit-and-run driver. She tried to call emergency services but the whole mobile network was down. She passed out and almost bled to death. Can you imagine what that must feel like?

Q: How are authorities getting smarter about shutting down the internet?

If we take the example of Ethiopia, where I am from – the first time they shut down the internet, it was like they didn't know what they were doing. They took us off the grid completely. That was extreme, and the economic cost was huge. So the next time they said, 'OK, we're just going to shut down mobile data' since shutting down broadband affects businesses more.

https://doi.org/10.14361/9783839449462-004 https://www.inlibra.com/de/agb - Open Access -

Governments also realize that they can limit a shutdown to a specific city or region now. This is very common in Pakistan and India. It's rare to see a national outage anywhere in the world. And when it's just a neighborhood or small city that is affected, it's harder to document.

Making shutdowns harder to document seems to be the main reason many governments are now opting to simply slow down the internet. It can be slow to the point where it can take a whole day to upload one photo to Twitter, but still be really hard to figure out whether someone is tampering with the internet. For instance in Togo or Cameroon, in countries that don't have the best infrastructure, it might just be that your bandwidth is having a bad day.

For the global community of technologists working to figure out how to spot, measure and analyze shutdowns, slowdowns are an especially big challenge. We can often use data from the Open Observatory of Network Interference (OONI) to confirm if a website is blocked or a shutdown took place, but it's much harder to verify if a deliberate slowdown is happening.

Q: What needs to happen to address these problems?

What keeps me up at night, are the shutdowns we are *not* able to document or understand why happen. Internet shutdowns and human rights violations go hand in hand. In some contexts when we're unable to document shutdowns, egregious human rights violations have happened. That's why we also need more tech companies to come on board with detection and documentation efforts. Google and Facebook are the first ones to know when the internet goes down because practically everyone uses their services. I feel they could be more open about sharing data about internet interruptions with us, and with people around the world.

► Further reading

- Open Observatory of Network Interference (OONI). https://ooni.torproject.org/
- Measurement Lab. https://www.measurementlab.net/
- Netblocks. https://netblocks.org/
- Oracle Internet Intelligence Map. https://map.internetintel.oracle.com/

https://doi.org/10.14361/9783839449462-004 https://www.inlibra.com/de/agb - Open Access -

Taxing social media in Africa

How much would you pay your government for a day's worth of WhatsApp messaging?

One after another, the governments of three countries in Africa, Uganda, Zambia and Benin have announced or imposed new taxes on mobile internet customers in 2018, leaving millions of Africans struggling to cover the costs of getting online. Only in Benin did protests result in quick abandonment of the tax plan.

Governments have imposed these levies to raise public revenues, and also argue that they are protecting the local telecommunications sector from competition from internet companies from abroad. But in practice, the (intended or unintended) consequence has been to push more people offline, increase barriers to getting online, and vastly limit freedom of expression and access to information – as well as access to goods and services that are now online.

Uganda imposed the first of these tax schemes in July 2018, forcing residents to pay a daily tax of 200 shillings ($0.053 USD) to use any one of 58 "over the top" (OTT) mobile communication apps. These include – but are not limited to – social media services like Facebook, Twitter, Instagram, and LinkedIn; instant messaging and voice communication apps WhatsApp, Snapchat, Skype; and dating sites like Tinder and Grindr.

The law in Uganda also placed a 1% tax on the use of mobile money, which is now the required method for airtime top up of SIM cards. With the average Ugandan already spending 15% of their monthly income for 1GB of broadband data, the new tax puts popular internet services out of reach for most people.

This is not just a matter of chatting with friends. As anyone in the region knows, WhatsApp in particular has become an essential platform for communication and information-sharing in Africa. Millions of people rely on WhatsApp groups to conduct business, communicate about local issues, read the news, and seek help in emergencies.

For many Ugandans, social media like Facebook and WhatsApp are a gateway to the rest of the internet. In an opinion piece for Global Voices, Ugandan blogger Pru Nyamishana wrote:

"The tax ignores a critical lack of digital literacy, particularly among poor Ugandans. When I interviewed women living in Bwaise, a slum in Kampala,

https://doi.org/10.14361/9783839449462-004 https://www.inlibra.com/de/agb - Open Access -

I learned that for them, WhatsApp and Facebook *are* the internet. These are the only platforms they know how to use. So with the new tax, they will be cut off altogether."

After the tax had been in effect for six months, the Uganda Communications Commission reported national internet usage rates had dropped by from 47.4 % to just 35 %.

On the heels of Uganda's initiative, Benin approved a similar tax in September 2018, targeting mobile messaging and 'Voice over IP' calls (like Skype). It drove up the cost of a single gigabyte of data by nearly 250 % but was repealed just days later, in the face of public protests.

The Zambian government announced a flat daily tax of 30 ngwees (US $ 0.03) on IP-based voice calls in August. Despite pushback from civil society and Zambia's Chamber of Commerce and Industry, government officials went ahead with the tax, arguing that it would raise public revenues, bolster local telecommunications enterprises, and help cover the cost of investments in infrastructure.

"Jobs such as call centre workers, talk time sellers, conventional call technicians will reduce drastically if more Zambians migrate to internet calls and create jobs in America and elsewhere," tweeted Dora Siliya, Zambia's Minister of Information and Broadcasting Services.

Although this reasoning rang hollow for many internet users, Siliya's argument is consistent with longstanding frustrations on the continent about foreign-owned OTT services that have captured markets for messaging and voice calls, changing the game for national telecom operators.

Countries in Africa are not alone in resenting how the data and advertising-driven business models of big tech bring few immediate benefits to local economies, while enriching technology companies in the United States. Google and Facebook are increasingly now also in the infrastructure game which will affect the power balance with telcos even further. Meanwhile, it's a fact that popular OTT services have helped fuel the uptake of mobile internet, and enabled local businesses to operate more efficiently. They have been critical to creating a virtuous cycle of record growth in internet use, network investments, and also telco profits.

In a region where governments are known for restricting free speech through censorship, internet shutdowns, surveillance and legal threats, civil society and independent media also view OTT tax schemes as an attack on free speech. In two other cases, this is clearly warranted.

https://doi.org/10.14361/9783839449462-004 https://www.inlibra.com/de/agb - Open Access -

In Tanzania, a so-called "blogger tax" was introduced in April 2018 alongside new restrictions for online content, in a clear effort to limit online expression. It requires Tanzanian bloggers, YouTube channel operators, and independent website owners to register and pay roughly $900 USD per year to publish online.

In August, the Mozambican government decreed that individual journalists and media outlets using both traditional and digital platforms now have to register and pay between $500 to $3,300 USD for an accreditation license that must be renewed every five years.

Taxes like these propagate the misconception that internet access and social media use are luxuries. But their outcomes – like the drop in internet use in Uganda – offer a compelling case study on the importance of establishing protections for net neutrality. What citizens have emphasized in protests, and what local researchers have also demonstrated, is that access to a truly open internet is a boon for local economies, education, public health and life in general.

► Further reading

- Offline and Out of Pocket: The Impact of the Social Media Tax in Uganda on Access, Usage, Income and Productivity, Pollicy, 2019. http://pollicy.org/wp-content/uploads/2019/01/Offline-and-Out-of-Pocket.pdf
- Taxed, throttled or thrown in jail: Africa's new internet paradigm, Global Voices, 2019. https://globalvoices.org/specialcoverage/taxed-throttled-or-thrown-in-jail-africas-new-internet-paradigm/
- Eastern Africa: New tax and licensing rules for social media threaten freedom of expression ARTICLE 19, 2018. https://www.article19.org/resources/eastern-africa-new-tax-and-licensing-rules-for-social-media-threaten-freedom-of-expression/
- Challenges and opportunities for advancing internet access in developing countries while upholding net neutrality, Nanjira Sambuli, 2016. https://www.researchgate.net/publication/302555638_Challenges_and_opportunities_for_advancing_Internet_access_in_developing_countries_while_upholding_net_neutrality

https://doi.org/10.14361/9783839449462-004 https://www.inlibra.com/de/agb - Open Access -

Tracking China's censorship of news on WeChat

In China today, it is nearly impossible to live life without WeChat. What began as a chat app, similar to WhatsApp or Facebook Messenger, has become an essential tool for everything from reading the news to paying for your morning beverage of choice.

After Facebook, WeChat is the most popular social media service in the world. The company now boasts more than 1,0825 billion individual users, along with more than 20 million registered public accounts. These public accounts are where many people in China get their everyday news and information. While many news outlets still maintain their own websites, virtually all media in the country also use WeChat as a publishing platform. Some publish their stories only to their WeChat pages, where followers can comment or discuss the stories of the day.

But of course, not all comments – or even media stories – are permitted to stay online. With its massive user base and powerful social influence, WeChat has become a major implementer of China's rigorous censorship regime. What is published on WeChat – and what the company censors at the state's behest – is a powerful indicator of government concerns about sensitive political issues.

With no transparency about what is censored or why, citizens and researchers are left to speculate and guess where the red lines are drawn.

A group of researchers at the University of Hong Kong have been working to track technical censorship on WeChat, using an innovative Web "scraping" system that captures millions of posts from the platform's most popular public accounts and makes them available to others in formats that can be visualized, mapped and understood in the context of time.

Summarizing the WeChatscope project in a story for Global Voices, Marcus Wang and Stella Fan explained their approach:

"Our team tracked more than 4,000 public accounts covering daily news through our computer program which visits (and periodically revisits) published articles and records the contents. When the system sees that a post has disappeared, it is detected as censored. A copy of the post is then restored in the database and made available for public access."

By the end of 2018, the group had identified roughly 11,000 posts that had been censored. These posts reflected some of the hottest and most controversial media stories and scandals of the year, ranging from the China-US

https://doi.org/10.14361/9783839449462-004 https://www.inlibra.com/de/agb - Open Access -

trade war, to tax fraud allegations against X-Men actress Fan Bingbing, to the #metoo movement at universities across China.

Explaining the context and possible reasons for censorship to a global audience is the subject of an article series on Global Voices, written in English and translated into multiple languages by volunteers. The stories describe in vivid detail how online speech in Chinese platforms can often initially be as vibrant, argumentative or controversial as elsewhere – despite censorship.

The WeChatscope project sheds light on what often feels like a black box of censorship policies and practices that are crafted and carried out by the Chinese government – and the companies required to comply with state demands. It also offers new possibilities for tech experts inside and outside the country to seek new ways to circumvent censorship in China.

► Further reading

- WeChatscope. http://wechatscope.jmsc.hku.hk/
- WeChatscope articles on Global Voices. https://globalvoices.org/author/wechatscope/
- What do Xi Jinping and Winnie the Pooh have in common? They're both flagged by Chinese censors, Shan Wang, Nieman Lab, 2018. https://www.niemanlab.org/2018/03/what-do-xi-jinping-and-winnie-the-pooh-have-in-common-theyre-both-flagged-by-chinese-censors/

Inside Germany's crackdown on hate speech

At the heart of the dilemma about what to do about the plague of hateful and harassing comments online, are questions of free speech, local laws and who should decide what can be said by whom.

Historically, internet companies have benefited from well established safe harbors from liability for the speech of their users, an approach that has helped enable the Web to become the creative and impactful environment it is today. However, hate speech and harassment have flourished online, and efforts by global platforms like Facebook, YouTube and Twitter to respond have been inconsistent and largely ineffective.

Germany (with a population of nearly 83 million people) recently thrust itself into the global spotlight on this question, implementing a law in 2018

https://doi.org/10.14361/9783839449462-004 https://www.inlibra.com/de/agb - Open Access -

intended to reduce hate speech and defamation online. The law introduces steep fines for popular social media companies if they do not take down manifestly unlawful content within 24 hours of a notification, and other unlawful content within up to seven days.

The Network Enforcement Act (NetzDG) was praised by some politicians as an important measure to curb hate speech and vehemently opposed by others. It was widely criticized by digital rights groups concerned about threats to free speech and overbroad takedowns. From abroad, it was observed with glee by governments who limit free speech. Russia, Venezuela and Kenya are among countries who quickly designed their own versions of the law.

In Germany, one year after implementation, the new law seems to be neither particularly effective at solving what it set out to do, nor as restrictive as many feared. However, without more insight into the kinds of notices that are being sent and the methods and guidelines platforms have adopted to handle them, it's difficult to assess the real impact.

NetzDG was designed to put the onus on companies to moderate content and remove it quickly. Germany's Federal Office of Justice can fine companies up to 50 million Euros ($56.3 million USD) if platforms fail to comply with valid removal requests by users or authorities. After the law was passed, Facebook and Twitter said they hired additional moderators in Germany to review content flagged as problematic by users or algorithms.

To comply with the law, Facebook, Google+, YouTube and Twitter each published reports in July 2018 and December 2018 detailing how they enabled users to file complaints and how they dealt with those complaints. So far, the number of content takedowns reported by platforms appears low compared to the number of complaints received.

Twitter, for example, said they received 256,462 complaints between July and December 2018 and took action on just 9%. Facebook said they saw 1,048 complaints and took down just 35.2% of reported content. What these complaints were about, or why so many were rejected, is unknown. Independent researchers have no access to raw data, and there is no standardized reporting process between platforms. The numbers are open to interpretation from every angle.

"If we want to better understand how companies make decisions about acceptable and unacceptable speech online, we need a more granular understanding of case-by-case determinations," wrote researchers from Germany's Alexander von Humboldt Institut für Internet und Gesellschaft in reac-

https://doi.org/10.14361/9783839449462-004 https://www.inlibra.com/de/agb - Open Access -

tion to the reports. They call for greater transparency and insight in order to understand what the effect of the law has been: "Who are the requesters for takedowns, and how strategic are their uses of reporting systems? How do flagging mechanisms affect user behavior?"

While most platform content rules are understood to be based on terms of service, community guidelines and other user policies, relatively little is communicated directly by platforms about how they enforce their own rules on prohibited content.

In Germany, an opportunity to come out of a contentious and politicized debate about harmful content with greater knowledge and better solutions has so far not materialized. Greater transparency around the sources of hateful and violent speech online, who reports it and how takedowns are approached by intermediaries would be an important step toward understanding how to foster a healthier internet for all.

► Further reading

- Removals of online hate speech in numbers; Kirsten Gollatz, Martin J. Riedl and Jens Pohlmann, Digital Society Blog, 2018. https://www.hiig.de/en/removals-of-online-hate-speech-numbers/
- Germany's NetzDG: A key test for combatting online hate, Olivia Knodt of the Counter-Extremism Project (CEP) and William Echikson of the Centre for European Policy Studies (CEPS), 2018. https://www.ceps.eu/system/files/RR%20No2018-09_Germany%27s%20NetzDG.pdf

Wikidata gives wings to open knowledge

How do voice assistants like Alexa and Siri know so much? How can search engines tell you the height of Mount Kilimanjaro (5,895 meters) so quickly and so accurately? Now more than ever, it is because they have access to more than 60 million open data records via Wikidata.

Wikidata is a project of Wikimedia, the non-profit organization that also runs the online encyclopedia Wikipedia. For six years, volunteer contributors to Wikidata have been structuring data so that it can be read and edited by both humans and machines.

https://doi.org/10.14361/9783839449462-004 https://www.inlibra.com/de/agb - Open Access -

This ensures that information can fly freely between the Web and other technology platforms. As more people interact with the internet, not just through the Web and websites like Wikipedia, but through speaking and listening to devices, this is becoming increasingly important.

Machines can understand Wikidata because it parses information you would normally read in a Wikipedia article into separate blocks. For example: "Paris is the capital city of France." "Paris has a population of 2,206,488." "Paris' coordinates are 48°51'23.68"N, 2°21'6.58"E."

By structuring this information and giving every entry a unique ID, Wikidata gives more than 5,000 websites, archives, libraries and databases a shared backbone: if you update one entry, other entries where the information is referenced will automatically be updated too, in every language.

Wikidata isn't the only initiative to organize, or try to organize, the Web's data. Similar projects have struggled due to its vastness. So what makes Wikidata successful? "Community is the biggest asset for Wikimedia," says Lydia Pintscher, Wikidata's Product Manager. "Without our partners and contributors, and the people who use the data, it wouldn't be there."

Indeed, Wikidata's community of tens of thousands of volunteer contributors have provided more than 850 million collective edits over the years.

Wikidata is also unique in that it is a completely open public domain resource: Their application of the Creative Commons CC0 Public Domain Dedication to all of Wikidata's data enables people and companies to use Wikidata freely and without copyright restrictions for whatever they like, from voice assistants to search engines.

For Wikidata, this open-access, no-citation approach means people benefiting from its information usually won't know where it comes from – or, that Wikidata depends on volunteers and donations to conduct its work, including updates and quality controls.

For big tech companies offering services on top of Wikidata and other Wikimedia properties, it ups the duty for them to help sustain the resource for everyone. "As companies draw on Wikipedia for knowledge – and as a bulwark against bad information – we believe they too have an opportunity to be generous," wrote Wikimedia's executive director, Katherine Maher, in an op-ed in WIRED in 2018 calling for companies to pay back to the community.

Companies including Google, Amazon and others have met this call to varying degrees (Amazon naming Wikipedia as part of the reason for Ama-

https://doi.org/10.14361/9783839449462-004 https://www.inlibra.com/de/agb - Open Access -

zon Alexa's success) but the vast majority of Wikimedia's resources come from donations by more than six million individuals who on average give $10 USD. In 2018, only 4 % of funding came from corporations.

For the health of the internet, open access to knowledge and information is essential. For institutions, companies, organisations and individuals with small and large data sets to share with the world, Wikidata is where it can really grow wings.

► Further reading

- A Brief Introduction to Wikidata, Björn Hartmann, Towards Data Science, April 2018. https://towardsdatascience.com/a-brief-introduction-to-wikidata-bb4e66395eb1?gi=d6b5aad15e2a
- Wikidata Tours. https://www.wikidata.org/wiki/Wikidata:Tours
- Amazon Owes Wikipedia Big Time, Slate, 2018. https://slate.com/technology/2018/10/amazon-echo-wikipedia-wikimedia-donation.html

"Deepfakes" are here, now what?

In a 2018 video, Barack Obama looked into the camera and warned: "We're entering an era in which our enemies can make it look like anyone is saying anything, at any point in time. Even if they would never say those things."

The video looks and sounds like Obama. But Obama never said those words.

The video is actually a *deepfake*: a photo, video or audio clip manipulated using AI to depict a person saying something that they have never said, or doing something they have never done.

The Obama deepfake was a project by filmmaker Jordan Peele and BuzzFeed CEO Jonah Peretti, intended to warn the public about misinformation online. Using free tools (and the help of editing experts) they superimposed Peele's voice and mouth over an existing video of Obama.

This kind of technology has long been available to Hollywood filmmakers. But in the last two years, it has taken a giant leap forward in accessibility and sophistication.

Deepfakes gained mass notoriety in 2018, with a wave of manipulated videos that used AI to put celebrities' faces onto porn actors' bodies. The term

https://doi.org/10.14361/9783839449462-004 https://www.inlibra.com/de/agb - Open Access -

deepfake itself comes from the handle of a Reddit user – Deepfakes – who made these kinds of videos and started the /r/deepfakes subreddit to share them.

The rise of deepfake porn prompted decisive responses from some platforms, several of which classified it as non-consensual pornography. The /r/deepfakes subreddit was banned in February 2018 for this reason.

But the name *deepfake* stuck. Possibly because it seems to make sense: 'deep' referring to 'deep learning' techniques used to create the media, and 'fake' referring to its artificial nature.

The technology is not only getting more accessible, but its applications are also expanding in multiple directions including producing full body deepfakes, creating real-time impersonations and seamlessly removing elements from videos. Concern is growing worldwide about the negative impacts that deepfakes could have on individuals, communities, and democracies.

The potential for harm is real. But Sam Gregory, Programme Director at the human rights organization WITNESS, says that instead of letting fear paralyze us, we need to focus on finding solutions. He published an extensive survey of solutions to malicious usage of deepfakes and synthetic media, based on conversations with experts in the field.

In the category of technical solutions, many platforms, researchers and startups are exploring using AI to detect and eliminate deepfakes. There are also new innovations in video forensics that aim to improve our ability to track the authenticity and provenance of images and videos, such as ProofMode and TruePic, which aim to help journalists and individuals validate and self-authenticate media.

While Gregory believes technical solutions are important, he says that they can't solve the problem alone. "It is vital to ask what communities might be excluded from technical solutions, and who has control over the data," he says. "If tools for tracking provenance become obligatory, they could be weaponized against individuals who can't access them or choose to remain anonymous."

Digital literacy is a critical solution that Gregory says is underexplored: "How do you get people to ask questions when an image looks flawless?" He says it's especially pressing to upskill people working with vulnerable groups and whose work could be negatively affected by deepfake technology, people like journalists and human rights advocates.

https://doi.org/10.14361/9783839449462-004 https://www.inlibra.com/de/agb - Open Access -

Many governments are grappling with how best to deal with online misinformation. But some activists and scholars caution against an outright ban of deepfake technology. They worry that if a law gives government officials the power to decide what is true or false, there is a risk that it might be used to censor unpopular or dissenting views.

Gregory also says civil society should develop a position on what role commercial platforms should play. "In many ways, platforms have the largest opportunity to detect deepfakes because they will have the largest body of training data. We should be clear now as civil society about what we want them to detect, and how we want them to inform the public, governments and key watchdog institutions."

Overall, Gregory cautions us to acknowledge the risks but resist the hype.

"It's good to not be apocalyptic about it, but to use this moment to have a rational discussion," he says, "The greatest harm of deepfakes may be to make people question everything."

► Further reading

- Deepfakes and Synthetic Media: Survey of Solutions against Malicious Usages, Sam Gregory, WITNESS, July 2018. https://blog.witness.org/2018/07/deepfakes-and-solutions/
- Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security, Robert Chesney and Danielle Keats Citron, California Law Review, 2019. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3213954
- Prepare, don't panic: dealing with deepfakes and other synthetic media, Sam Gregory, 2019. https://www.journalismfestival.com/programme/2019/prepare-dont-panic-dealing-with-deep-fakes-and-other-synthetic-media

► Further listening

- Breaking News, Simon Adler, RadioLab podcast, July 2017

https://doi.org/10.14361/9783839449462-004 https://www.inlibra.com/de/agb - Open Access -