← back home

Against an Increasingly User-Hostile Web

We're quietly replacing an open web that connects and empowers with one that restricts and commoditizes people. We need to stop it.

- Parimal Satyal, 2 november 2017.
See also: my Paris Web 2020 talk based on this article and the Hacker News discussion threads from 2017 and from 2020.

I quit Facebook seven months ago.

Despite its undeniable value, I think Facebook is at odds with the open web that I love and defend. This essay is my attempt to explain not only why I quit Facebook but why I believe we're slowly replacing a web that empowers with one that restricts and commoditizes people. And why we should, at the very least, stop and think about the consequences of that shift.

The Web: Backstory

*(If you want, you can skip the backstory and **jump directly to the table of contents**).*

I love the web.

I don't mean that in the way that someone might say that they love pizza. For many of us in the early 2000s, the web was magical. You connected a phone line to your computer, let it make a funny noise and suddenly you had access to a seemingly-unending repository of thoughts and ideas from people around the world.

It might not seem like much now, but what that noise represented was the stuff of science fiction at the time: near-instantaneous communication at a planetary scale. It was a big deal.

I was an average student at school. Despite well-meaning and often wonderful teachers, I didn't thrive much in a school system that valued test performance and fact-retention over genuine curiosity. Had it not been for the web, I might have convinced myself that I was a poor learner; instead, I realized that learning is one of my great passions in life.

What remains of my fan site for German powermetal band Gamma Ray from 2001, archived thanks to the wonderful folks over at Archive.org

I was 11 when I set up my first website. Growing up in Nepal, this *was* magical. Almost everything I love today—design, aviation, cosmology, metal music, computation, foreign languages, philosophy—I discovered through the many pages that found their way to my web browser. All I needed were curiosity, a

phone line and that strange little electrical song. And good old Netscape Navigator.

Netscape Navigator 4.04, source: A Visual Browser History, from Netscape 4 to Mozilla Firefox

The web enabled that. It's one of humanity's greatest inventions. And now, we the architects of the modern web—web designers, UX designers, developers, creative directors, social media managers, data scientists, product managers, start-up people, strategists—are destroying it.

We're very good at talking about *immersive experiences*, *personalized content*, *growth hacking*, *responsive strategy*, *user centered design*, *social media activation*, *retargeting*, *CMS* and *user experience*. But behind all this jargon lurks the uncomfortable idea that we might be accomplices in the destruction of a platform that was meant to empower and bring people together; the possibility that we are instead building a machine that surveils, subverts, manipulates, overwhelms and exploits people.

It all comes down a simple but very dangerous shift: the major websites of today's web are not built for the visitor, but as means of *using* her. Our visitor has become a data point, a customer profile, a potential lead — a proverbial fly in the spider's web. In the guise of *user-centered design*, we're building an increasingly *user-hostile* web.

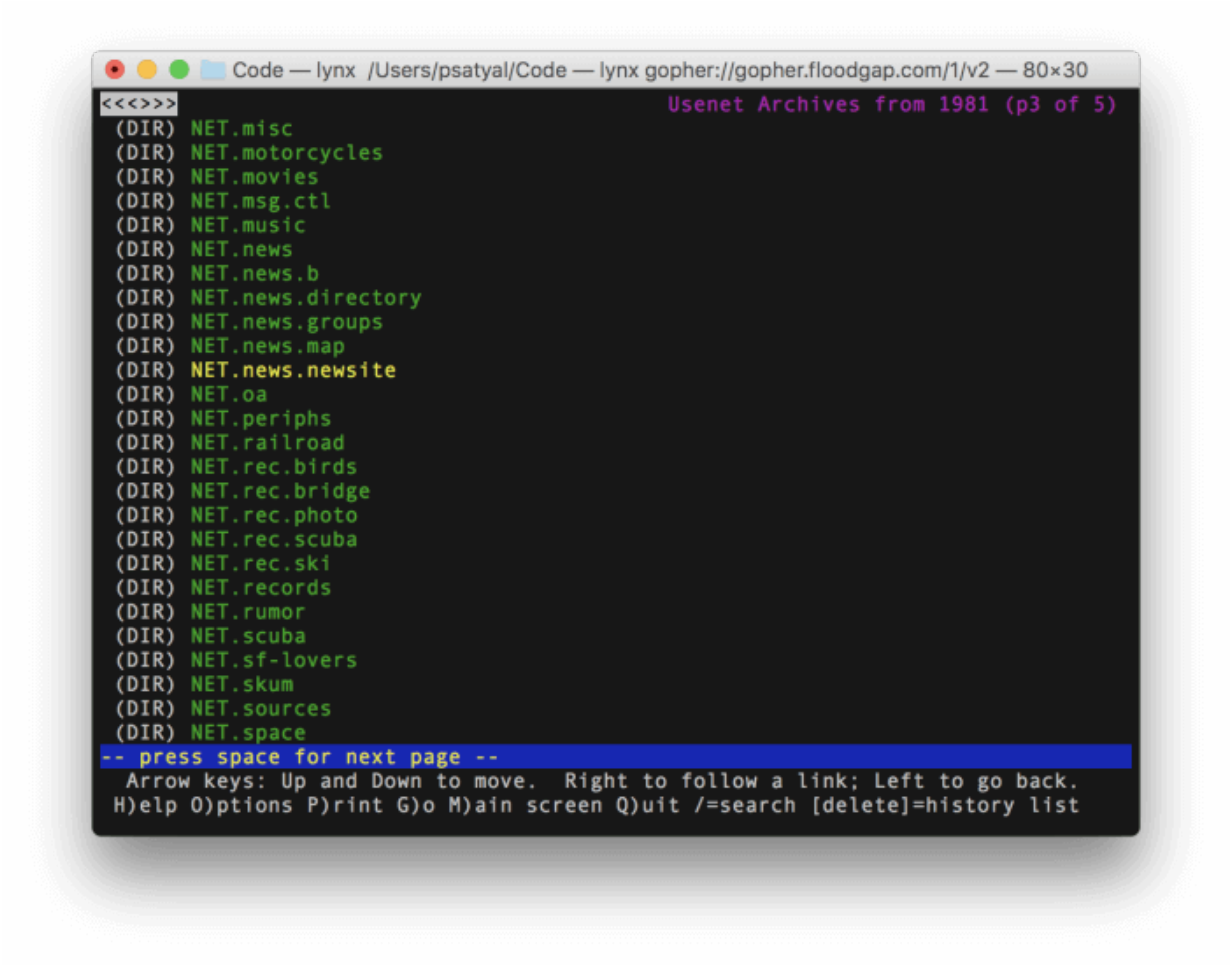
If you work in the design/communication industry, consider this essay introspective soul-searching by one of your own. If you're a regular web user, consider this an appeal to demand a better web, one that respects you instead of abusing and exploiting you.

Note: The entire essay is rather long so feel free to skip to individual parts:

The Web was Born Open

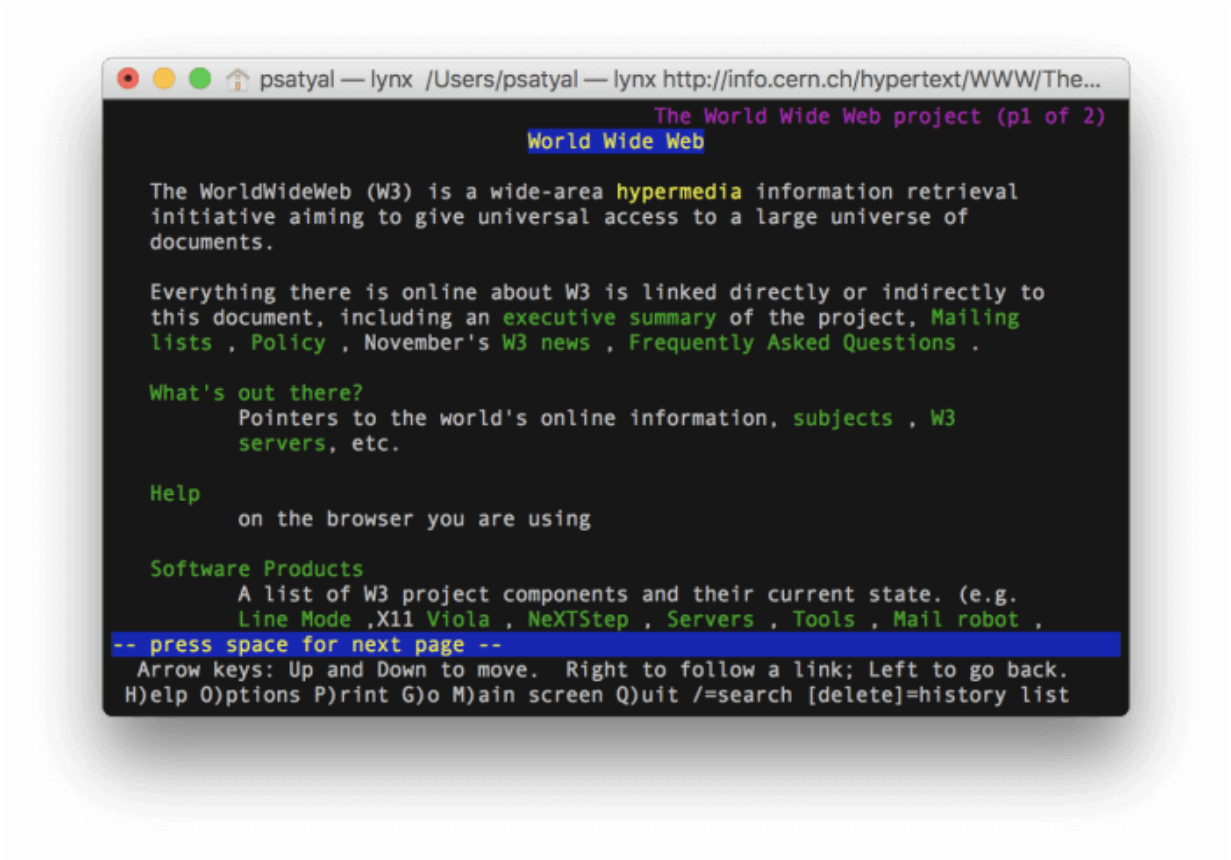
It all began in the early 90s.

The Internet—the physical network that allowed computers around the world to communicate—was already in place but it remained inaccessible to most people. You had to know how to use a local client to connect to a remote FTP, Usenet, Gopher or an email server. This was before the days of ubiquitous graphical user interfaces so you had to type funny commands into a terminal, one of those black screens with green text that that hackers supposedly use to do *Bad Things*.



Usenet Archives from 1981 on gopher server Quux.org, accessed 31 October 2017 via lynx

Meanwhile, Tim Berners-Lee was working as an independent contractor at CERN in Geneva. Frustrated with how difficult it was to find, organize and update technical documentation, he proposed a solution that involved "global computer networked information system" that "presented users with a web of interlinked documents", called *Mesh*. Pretty soon it became apparent that WWW—World Wide Web, as it came to be known—could do more than just link technical documents.



The world's first website, accessed 31 October 2017 via lynx

On April 30 1993, CERN made a bold decision. It decided to release WWW into the public domain. It renounced all intellectual property rights and essentially invited anyone at all, anywhere in the world, to play with it. Later, the director of CERN who approved the decision said that he was inspired by Richard Stallman's vision of free, open software.

Had CERN decided otherwise and patented the technology to then license it for money, the web would arguably not have taken off the way it did. It might have died out like the Minitel did in France. The web as we know it was born of a vision to create an open system that brought people and ideas together, with documents that "may reside on any computer supported by that web".

Advances in the hyper-text transfer protocol (HTTP), network infrastructure, web browsers and standards, consumer Internet access, accessible hosting and blogging platforms led to a massive democratization and adoption of the web.

Soon, anyone could put a document on the web and any document could link

to any other. It created a completely open platform where a writer in Nepal could freely share her ideas with a dancer in Denmark. A climate science student in Nairobi could access data from the McMurdo weather station in Antarctica. You could start reading about logical fallacies and end up on a website about optical illusions. Read about the history of time-keeping and end up learning about Einstein's special theory of relativity. All interests were catered to. Information could truly be free: transverse borders, cultures and politics.

That is the web at its best.

My own journey from designing that first website as an 11-year old "webmaster" in Nepal to writing this article as a UX Consultant in France has its origin in that 1993 decision by CERN.

The Modern Web (of Deception)

The modern web is different.

It's naturally different from a technological standpoint: we have faster connections, better browser standards, tighter security and new media formats. But it is also different in the values it espouses. Today, we are so far from that initial vision of linking documents to share knowledge that it's hard to simply browse the web for information without constantly being asked to buy something, like something, follow someone, share the page on Facebook or sign up to some newsletter. All the while being tracked and profiled.

Almost every website you go to today reports your activities to third parties that you most likely neither know nor trust. They record where you come from, what pages you visit, how long you stay on each, where you click and where you go next. In fact, since so many websites report to the same third parties, these companies can essentially have your web history on file as you go from link-to-link, website to website. Like an omnipotent eye embedded on Sir Berners-Lee's global system of interlinked documents, noting down everything you do and reporting to private entities who then sell this information for profit.

These companies build profiles, anonymous at first, with your interests and navigational behavior. These profiles can then get increasingly personal: they might include your email addresses, home address, income, educational history, political affiliation, information on your family. Over time, they can cross-reference all this information with your location data to figure out where you work, which restaurants you go to, where your gym is. Recently, we even learned that Google was able to associate your offline purchases with your online ad viewing history (albeit *anonymously*, it would appear). Once they have that, they can look into your behavior and psychology: what kind of ads do you tend to click on? What kind of messages resonate most with you? What are the best strategies to influence your opinion?

Screenshot of Mr. Alexander Nix presenting the work of Cambridge Analytica, video The Power of Big Data and Psychographics on Youtube

The Leave campaign responsible for Brexit in the United Kingdom and Donald Trump's 2016 presidential campaign both bought the services of a certain Cambridge Analytica, a company that boasts a gigantic database containing personal details amounting to "close to four or five thousand data points on every adult in the United States" (their own words). The goal? Craft hyper-personalized messages to change voting behavior based on your individual personalities, and by extension, your attitudes, opinions and fears. So if you are identified as a dad of three young kids in rural Texas, the message is nuanced to suggest that only a certain candidate will be able to protect your

family against real or imagined threats. If you are identified as a patriot who's previously posted comments about gun rights and the second amendment, it might be about crime rates and how the opposition is trying to take your constitutional rights away from you.

> You become a manipulable data point at the mercy of big corporations who sell their ability to manipulate you based on the data you volunteer.

This is the equivalent of someone following you in real life as you go about your everyday business, like a private eye who notes down with whom you meet, what you talk about, what you spend time looking at in stores. A private eye who takes notes and then sells it to the highest bidder. But you got to enter the store for free, so you should be so glad. The stores might also justify it. *"Sure it's a bit invasive, but we'll be able to give you better recommendations if we know what you like"*.

But how do they get all this personal information — where you live, who your friends are, what your religion and ethnicity are, where you were last night, what you bought on Monday? Most of it you volunteer yourself on social platforms like Facebook, Twitter and Instagram. The little share buttons you see on websites aren't just there to make it easy for you to post a link to Facebook; they also allow Facebook to be present and gather information about you from pretty much any website.

But how can you know that any of this is true?

Track the Trackers: An Experiment

Perhaps you think I'm being a tad too dramatic.

In your defense, all of this does sound like some dystopian fantasy. But I'm not that great a fiction writer quite yet. Let me illustrate my point with a little experiment. We'll pick a major website that you might visit regularly and identify third parties it shares your information with.

We'll need a few things:

- a test website
- Webbkoll, a web privacy check tool by Dataskydd.net, a Swedish association for data protection and privacy (of which I'm a proud member) and
- A web inspector

Let's take an article that was published around the time I first started working on this article (which is last year; I'm a slow writer): **Astronomie : la sonde Juno s'est mise en orbite autour de Jupiter** (*Astronomy: space probe Juno put in orbit around Jupiter*).

Le Monde article Astronomie : la sonde Juno s'est mise en orbite autour de Jupiter

If you run this URL through Dataskydd's Webbkoll and a web inspector tool (I used Chromium's web inspector), you learn a few interesting things: the page is **3.1 MB** in size, makes about **460 HTTP requests** of which **430 are third-party requests** (outside of its parent domain) and takes **20 seconds** to fully load on a fast 3G connection (from Paris, France).

It also stores **100 cookies** (these are little pieces of text stored on your computer by websites other than lemonde.fr; cookies are normally used to save session information but can also be used to identify and track you) and contacts **118 third-parties**. And if all this weren't enough, your connection to LeMonde and the majority of third-party connections are over **unsecure HTTP** protocol (instead of the more secure HTTPS, which should be a basic requirement).

That's a lot of big numbers for an article of 1500 words, three images and one video.

Now let's look at some of the third parties that the page connects to when you load it:

- Weborama: advertising platform for analytics, digital marketing and behavioral targeting
- Visual Revenue: predictive analytics platform
- AppNexus: multimedia content monetization service
- Outbrain: "online advertiser specializing in presenting sponsored website links" (Wikipedia)
- Facebook: a social network and micro-targeted advertising platform
- Cedexis: a multi-CDN application delivery platform

Note: In an earlier version of the article, I had mistakenly identified Cedexis as an "ad-delivery platform", which it is not. My apologies to Cedexis for the error.

Some of these are simply tools to manage content delivery but many are advertising or content monetization platforms. Companies like Weborama make money by selling information about you. When people say, "you're the product," it isn't just some analogy, it accurately reflects the business propositions of many such companies.

What's surprising is that the bulk of the information transferred between LeMonde and you doesn't even concern the actual article. If you were to isolate the actual content—the words, images and video—and put it in an HTML file, it would weigh considerably less than 3.1 MB and would make a lot

fewer requests.

If fact, I did just that and made three versions :

- Version A: With the original text (including comments, images and video)
- Version B: With the original text (including comments, images) but no video
- Version C: With just the original text (including comments), no images or video

Some numbers:

	Original (LeMonde.fr)	Version A	Version B	Version C
Page Size	3,1 MB	1 MB *(32%)*	183 KB *(5,8%)*	17 KB *(0,54%)*
Load Time	20,9 s	4,6 s *(19,4%)*	2,8 s *(9,6%)*	662 ms *(3,2%)*
Requests (total)	459	108 *(23,5%)*	5 *(1%)*	1 *(0,2%)*
Requests (third-party)	436	64 *(14,7%)*	4 *(0,9%)*	0
Third Parties Contacted	118	17 *(14,4%)*	2 *(11,8%)*	0
Cookies (total)	100	16 *(16%)*	0	0
Cookies (third-party)	73	16 *(21,9%)*	0	0
Text *(% of Page Size)*	0,5 %	1,7 %	9,5 %	100 %
Text + Images *(% of Page Size)*	5,8 %	17,9 %	100 %	
Text + Images + Video *(% of Page Size)*	32,3 %	100 %		

Note: Data on the number of requests (first- and third-party) and cookies (first- and third-party) comes from Dataskydd Webbkoll. The rest of the data comes from Chromium's built-in web inspector. All connections were made from Paris, France with cacheing disabled and the bandwidth throttled to simulate a "fast 3G" connection. You can run these numbers yourself; they should vary only nominally depending on where you are. If you find errors, please let me know.

Those are some very interesting figures. Some observations:

- The actual article (text and three images, *version B*) **makes up less than 6% of the total size of the page** on LeMonde.fr. This means that **94% of the data transferred between you and LeMonde.fr has nothing to do with the article**.
- What about the video, you ask? Before you even play it, **that one video adds over a 100 requests (60 of which are to 15 additional third parties) and 16 third-party cookies**. It also adds over 800 KB of data. Again, this is before you even decide to play the video. The video might be related to the content, but it's doing a lot more than that.
- Even compared to the version with the video *(Version A)*, **the LeMonde article makes about 450 additional third party requests, of which 370 are to about 100 additional third parties, storing 100 additional cookies (55 of which are third party cookies)**. It also adds over 2 MB to the page. All that is data that has nothing do with and is completely unnecessary to load the article you're reading.
- The text + image version *(Version B)* is **able to load the entire text and the 3 images with only 5 requests and no cookies whatsoever**. Adding a video should reasonably add one or two more requests and maybe one cookie, not 450 requests and 100 cookies, the majority of which are on behalf of companies you neither know nor trust, including those who track and sell your data for profit.
- The Le Monde page will **continue to periodically transfer data and make additional requests even after it has completely loaded** and as you scroll and interact with the page. If you monitor network traffic, a lot of this data is going to third-party tracking scripts. For example, a request is made to Xiti.com (a web analytics company) every few seconds.
- If you don't use a content blocker, you will notice that **in just a matter of minutes, over 30 MB of data will be transfered between your browser and the 100+ third parties**. The number of requests will go into the thousands. This will continue to rise as long as you leave your browser open.

Essentially, this means that about 94% of the data being transferred and 99% of the requests being made have nothing to do with the article itself. Le Monde might principally be a newspaper in its printed version, but the online version is an invasive, insecure advertising platform with good content (in that order).

If you're curious, try using **Webbkoll** on other websites you visit to see how

privacy-friendly and respectful these websites are. We'll get into how to protect yourself from these third-party trackers later on in the article.

All this might not be illegal (although there's some doubt, especially now that in the context of up the upcoming European General Regulation on Data Protection), but it is rather disrespectful towards the user. Not only are these websites breaking my trust—when I visit your website, I entered into contact with you, not 80 other websites—but they are loading content from websites neither know nor trust. Some of which have been know to spread malware.

> Using an ad/content-blocker isn't cheating the system; it's taking very basic precautions that websites like Le Monde can't be bothered to take to protect you. For me, it's a basic necessity in the modern web.

If you're reading this and are wondering what to do to protect yourself, skip ahead to the The Way Forward section.

If you run a website and you put official share buttons on your website, use intrusive analytics platforms, serve ads through a third-party ad network or use pervasive cookies to share and sell data on your users, you're contributing to a user-hostile web. You're using free and open-source tools created by thousands of collaborators around the world, over an open web and in the spirit of sharing, to subvert users.

Gated Communities

One of the most impressive things about the Internet (and consequently also the web) is that it is decentralized. No central authority gets to decide which page is more important than others and you don't have to play by anyone else's terms to publish and read what you want. There isn't anything like a *main server* that stores the code that runs the Internet; it's just a protocol on a physical backbone (of undersea cables).

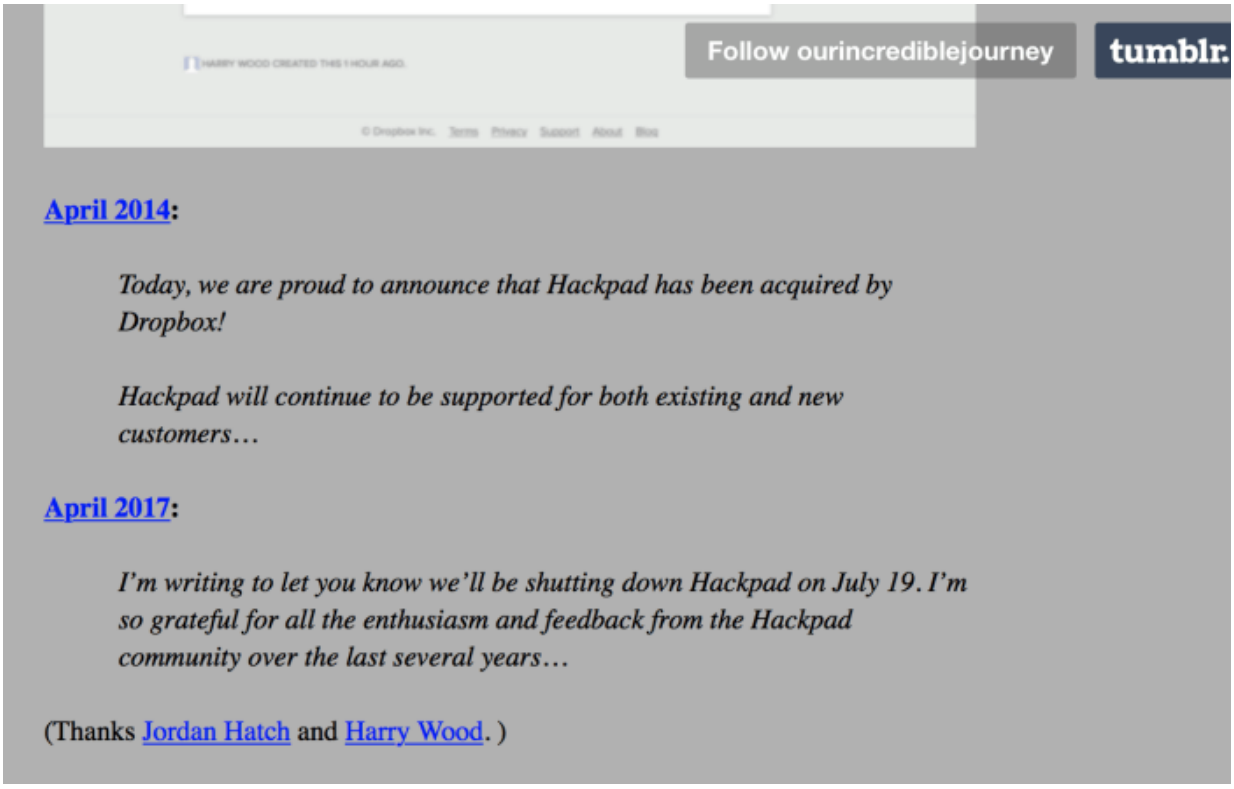
You could buy a Raspberry Pi Zero today for less than 10€, connect it to the Internet, set up a chat server on it, give it a public address and the world

would be able to connect to it and talk to one other. Sure, it might not perform too well and no one might actually use it, but it is technically possible.

But most of the time we spend on the web today is no longer on the open Internet - it's on private services like Facebook, Twitter and LinkedIn. While Facebook provides a valuable service, it is also a for-profit, company. Their source of revenue is advertising. It is the epitome of centralized.

Francisco Goya's *The Naked Maja* (1800)

Try posting a picture of the Francisco de Goya's "The Naked Maja" or your naked breasts (if you're a woman) on Facebook; it'll almost certainly be removed. It's against their terms of use. To use their platform, you have to agree to whatever conditions they set, however absurd. If you replace the open web with Facebook, you're giving up your right to publish and share on your terms. The data that you post there does not belong to you; you're putting it in a closed system. If one day Facebook decides to shut down—unlikely as that might seem today—your data goes with it. Sure, you might be able to download parts of it, but then what?



Tumblr Blog Our Incredible Journey, *"cataloging the thrilling opportunities start-ups are offered when their incredible journey continues by being bought by an exciting company. However, as a user of the start-up's service, your own incredible journey must end, because all of your photos and writing and checkins and messages and relationships must now be deleted"*.

This works because they know you'll agree to it. You'll say you don't have a choice, because your friends are all there—the infamous "network effect". This is Facebook's currency, its source of strength but also a crucial dependency.

> And this is what we often fail to realize: without its users—without you—Facebook would be nothing. But without Facebook, you would only be inconvenienced. Facebook needs you more than you need it.

And they do their best to keep you on their website as long as possible. Your attention is worth a lot to a lot of companies who are convinced that traditional advertising is dead and that micro-targeted campaigns work better. (And they mostly do, from their point of view). This drives them to come up with absurd techniques to create addiction: wish your friend happy birthday, wish your colleague a happy work anniversary (who does that?),

here's a video we made about you, three friends are going to an event near you, continue watching the video you started even as you scroll, be the first to comment, react to this photo, tell everyone what you're up to. The longer you stay, the more information you give, the more valuable your profile—and the platform—is to advertisers.

I'm not saying that what Facebook is doing is entirely unethical. It has to make money to make up for the resources it employs to keep the website running and it does so by advertising. Every time you choose to use a free service like Instagram, LinkedIn, Gmail or Snapchat, you are paying for the convenience with your eyes, your data and your attention. There's nothing inherently wrong as long you as you understand and consent to this exchange of value. But do you? Does your daughter? Your dad?

What I'm against is the centralization of services; Facebook and Google are virtually everywhere today. Through share buttons, free services, mobile applications, login gateways and analytics, they are able to be present on virtually every website you visit. This gives them immense power and control. They get to unilaterally make decisions that affect our collective behavior, our expectations and our well-being. You're either *with them or out*. Well, I chose out.

You see, the web wasn't meant to be a gated community. It's actually pretty simple.

A web server, a public address and an HTML file are all that you need to share your thoughts (or indeed, art, sound or software) with anyone in the world. No authority from which to seek approval, no editorial board, no publisher. No content policy, no dependence on a third party startup that might fold in three years to begin **a new adventure**.

A website on Doom level design on Geocities from 1999, accessed October 31, 2017 via Archive.org

That's what the web makes possible. It's friendship over hyperlink, knowledge over the network, romance over HTTP.

In fact, the browser you're reading this on (Chrome, Firefox, lynx, whatever), the web server that's hosting this website (Nginx), the operating system that this server runs on (Ubuntu), the programming tools used to make it all work (python, gcc, node.js...) — all of these things were created collectively by contributors all around the world, brought together by HTTP. And given away for free in the spirit of sharing.

The web is open by design and built to empower people. This is the web we're breaking and replacing with one that subverts, manipulates and creates new needs and addiction.

The Way Forward

If you want to protect yourself (as a user) from predatory web marketing companies and defend the open web, there a few things you can do today at an individual level.

If you're a web professional (a designer, UX consultant, strategist, programmer…), there are a number of considerations for better respecting your users and protecting their privacy (and your integrity).

Here's a basic list:

For end users (you, dear reader)

- If you use Chrome as your main browser, consider switching to the open-source version called Chromium. Better yet, switch to Mozilla Firefox, developed by the not-for-profit Mozilla Foundation that has a solid record of defending your privacy. Consider minimalist browsers like Min (and choose to block all ads, trackers and scripts) to browse news websites.
- Install a content/ad blocker for your browser: I recommend uBlock Origin (available for Firefox, Chrome and Safari on most platforms). You can also complement this with the Electronic Frontier Foundation's Privacy Badger tool that protects you from invasive ads and third-party tracking.
- Install HTTPS Everywhere for your browser; this forces your information through secure, encrypted channels (HTTPS vs HTTP one) if possible. It can also be configured to only allow connections to HTTPS websites.
- Think about how much information/details you provide to social media platforms like Facebook, LinkedIn, Twitter and Instagram. They already have quite a lot (including the ability to recognize you by name on photographs), but what other information are you volunteering? Where you are, whom you're with, information about your friends?
- Consider quitting social networks, especially Facebook (but download your data first!). What would you miss the most? Are there alternatives?
- Consider alternatives to free services provided by the likes of Google and Facebook. Today, if both of these companies shut down (or implement policies I don't like), I would mostly be fine because my contact with them is limited. I use DuckDuckGo and Startpage for search (free); FastMail for email and calendar (less than 40€ a year) ; HERE WeGo for maps (free); Signal, email and IRC for messaging (free, along with iMessage, Whatsapp and Twitter); Digital Ocean for web hosting (about 5€ per month).
- Pay for services and content that you like, if you are able. If you like reading The Guardian, for example, consider subscribing. If your favourite YouTube channel is on Patreon, consider pledging a small amount per video. If you like

services like Pinboard.in that charge in return for a useful service, buy it. There's mutual respect when both the user and the service provider know what basic service they are buying/selling.

- At the very least, consider that the platforms you use need you more than you need them. You have power over them (unfortunately, in numbers) and they know it. If enough people care about privacy and respect for their data and time, platforms will have to adapt to stay relevant.

For web professionals (you, fellow industry colleague)

- Consider not putting share buttons everywhere. They're visual noise and make third party connections every time the page is loaded (adding to load time). If you have to, create your own instead of using ones provided by Facebook and co. (so that a click is needed before a request is made to their servers)
- Support HTTPS. It's super easy (and free!) with Let's Encrypt so you don't have an excuse to not respect your users' privacy
- Think about accessibility also in terms of page size, load times and tech requirements: will your website work without Javascript? What percentage of the total weight of your page is actual information? How many third party requests are you making? How long would it take to load on a 56.6k dial-up or on EDGE? How does it render for speech readers? Can it be read via a text-based browser? (It's a fun experiment; try visiting your website with a text-based browser like lynx or Links).
- Refuse client requests to implement hyper-invasive technologies like canvas fingerprinting.
- Consider replacing Google Analytics with a more privacy-respecting analytics software like Piwik. Even better if you can host it yourself!
- Minimize third-party dependencies like Google Fonts (you can self-host them instead).
- Avoid ad networks (like the plague!) if possible. Serve your own ads by selling ad space the old school way if you're able. If not, explore privacy-respecting methods of serving ads, including developments powered by the blockchain (like the Basic Attention Token).
- Respect Do Not Track.
- Carefully consider the benefits of hyper personalisation and retargeting. The benefits are debatable but the long term consequences might be disastrous.

Ask yourself: would you be okay with a company collecting as much data (as you seek to collect) on your teenage daughter, your nephew in college, your husband or your grand-mother?

- Consider business models where you actually respect your clients and your website visitors instead of using them. If you can't be honest about your business model with your client, maybe you need to ask questions.

Thoughts and feedback

It all comes down to one simple question: **what do we want the web to be?**

> Do we want the web to be open, accessible, empowering and collaborative? Free, in the spirit of CERN's decision in 1993 or the open source tools it's built on? Or do we want it to be just another means of endless consumption, where people become eyeballs, targets and profiles? Where companies use your data to control your behaviour and which enables a surveillance society—what do we want?

For me, the choice is clear. And it's something worth fighting for.

I hope this article has been interesting. If you have thoughts—you agree, disagree, have reservations, other ideas or a suggestion—I'd love to hear them! This **article is on GitHub**; if you'd like you can send a pull request with edit suggestions (like **Anders** and **many others** did, thank you!). You can also get in touch via email (*userhostileweb—at—neustadt.fr*) or, if you're on **Hacker News** or Reddit, share your thoughts there.

—

← back home