

As the Pirates Become CEOs: The Closing of the Open Internet

Zeynep Tufekci

Abstract: The early Internet witnessed the flourishing of a digitally networked public sphere in which many people, including dissidents who had little to no access to mass media, found a voice as well as a place to connect with one another. As the Internet matures, its initial decentralized form has been increasingly replaced by a small number of ad-financed platforms, such as Facebook and Google, which structure the online experience of billions of people. These platforms often design, control, influence, and “optimize” the user experience according to their own internal values and priorities, sometimes using emergent methods such as algorithmic filtering and computational inference of private traits from computational social science. The shift to a small number of controlling platforms stems from a variety of dynamics, including network effects and the attractions of easier-to-use, closed platforms. This article considers these developments and their consequences for the vitality of the public sphere.

ZEYNEP TUFEKCI is Assistant Professor at the School of Information and Library Science, with an affiliate appointment in the Department of Sociology, at the University of North Carolina, Chapel Hill. She is the author of *Beautiful Tear Gas: The Ecstatic, Fragile Politics of Networked Protests in the 21st Century* (forthcoming 2016) and coeditor of *Inequity in the Technopolis: Race, Class, Gender, and the Digital Divide in Austin* (with Joseph Straubhaar, Jeremiah Spence, and Roberta G. Lentz, 2012). She is also a contributing opinion writer for *The New York Times*.

I traveled to Cairo in the spring of 2011, a few months after the fall of President Hosni Mubarak. Egypt was unsettled but jubilant, and the rest of the Middle East had not yet fallen into war or renewed authoritarianism. One of the Egyptians I interviewed was a blogging pioneer whom I will call Hani.¹ In the early 2000s, Hani had been among the first to take advantage of the burst of freedom experienced by Egyptians before the authorities fully caught on to the Internet’s revolutionary potential. Many bloggers made it through the Mubarak era largely unscathed because the government could not keep up with or fully understand the new medium. Unfortunately, Hani had caught the attention of the government; he was tried and sentenced to years in prison for the crime of insulting Mubarak. Throughout his imprisonment, he remained defiant. He was released in November 2010, just months before a Facebook page would spark a revolution that would dramatically change the country, the region, and the world.

Before going to jail, Hani felt that his blog had been a bustling crossroads of discussion. His voice reached

© 2016 by Zeynep Tufekci
doi:10.1162/DAED_a_00366

farther than he had ever thought possible. After his multiyear involuntary hiatus, he resumed blogging, he told me, with enthusiasm and excitement. But when he came out of jail in late 2010, he found that his blog, and much of the Egyptian blogosphere, had become a comparative wasteland.

“Where is everybody?” Hani answered himself: “They’re on Facebook.”

At the time I interviewed him, that did not seem like such a bad development. Just a few months earlier, a Facebook group titled “We are All Khaled Saed” – named after a young Egyptian man who had been tortured and killed by the police – had become the organizational core of the revolution. The page was created in June 2010, a few days after Saed’s death became public knowledge. The (then-anonymous) administrator of the page was Wael Ghonim, a Google employee and early adopter of the Internet in the region. Ghonim had foreseen Facebook’s potential to reach large numbers of ordinary people: in just one month, his page gathered more than one hundred thousand readers, and ordinary Egyptians began using it to engage in political discussion.² In later interviews, some of those who participated on the page told me that they felt jubilant and liberated to be finally speaking about politics with other Egyptians.

After the Tunisian revolution of early 2011, the “We are All Khaled Said” page became a hotbed of conversation for Egyptians who longed for a similar upheaval. After much discussion, including polls asking the page’s readers what they thought should be done, Wael Ghonim created an event titled “The Revolution,” scheduled for January 25, 2011, which was already a traditional day of protest in Egypt.³ Hundreds of thousands of Egyptians accepted an “evite” to “The Revolution,” displaying their dissent openly, many perhaps for the first time, to their online social networks.

Emboldened by the outpouring of dissent, thousands of people assembled in Tahrir

Square on January 25, 2011. One year prior, only about one hundred protesters met in Tahrir Square, where they were surrounded and outnumbered by the police. But this year, the protest quickly swelled to include hundreds of thousands of Egyptians who occupied the square until Mubarak stepped down. To many activists I talked with, Facebook’s reach felt empowering. A survey of Tahrir protesters confirmed that social media had been essential to the early turnout that had triggered the avalanche of dissent.⁴ Egyptian use of Facebook continued to grow, and it became plainly obvious that Facebook had become a major player in the civic sphere. Even the new military council that replaced Mubarak launched a Facebook page.

But what did it mean for Facebook, a corporate platform, to become so central to the political life of the country? That was less clear.

With the advent of social media platforms in the mid-2000s, the “networked public sphere” – the burgeoning civic space online⁵ that had been developed mostly through blogs – expanded greatly, but with a simultaneous shift to commercial spaces.⁶ Many scholars and civic activists worried about how “sovereigns of cyberspace,” as Internet-freedom advocate, journalist, and author Rebecca MacKinnon called these online platforms, would wield their power.⁷ Would they censor and restrict freedoms to serve advertisers or governments with whom they were trying to curry favor? Would they turn over user information to repressive regimes? MacKinnon was prescient in identifying the core problem: the growth of privately owned spaces that functioned as public commons. Over time, the threats posed by this relationship may exceed even our earlier concerns about censorship.

Driven by structural dynamics and corporate motivations, as well as by characteristics of the Internet, these new social platforms are remaking the Internet in a way

that imperils the open architecture of the early Web that felt so intoxicatingly empowering to many of its users. The consequences are profound. This article examines where we are now, and then briefly traces the dynamics that have led us here.

In 2015, Hossein Derakhshan – who has been called the “grandfather” of the Iranian blogosphere – left prison after serving six years of a nineteen-year sentence for blogging, including long stretches of solitary confinement. But prison did not break him; instead, he says, what nearly broke his heart was what he found online when he started blogging again.⁸

After being released, Derakhshan learned that he needed to adapt to the new digital environment and use the new commercial social networks. Up for innovation and change, he created a Facebook account and posted a link to his blog. To his dismay, his post disappeared after just a few “likes.” Likes are the main currency in Facebook’s all-important algorithm that decides which posts to display to other users, and which to hide. In the new world of social media, posts like Derakhshan’s could disappear without being seen by more than a handful of people. Derakhshan was despondent about trying to learn the ropes of this new world. But he soon realized that his personal grasp of the platform was not the only missing ingredient.

The new platforms were strangling access to the hyperlink, directing users to content within their walls and regulating access to the outside Web in very specific ways. Content like his, which was hosted outside of Facebook’s territory, did not stand a chance.

Derakhshan wrote the essay “The Web We Have to Save” about his new experience of being online:

Nearly every social network now treats a link as just the same as it treats any other object – the same as a photo, or a piece of text – instead

of seeing it as a way to make that text richer. You’re encouraged to post one single hyperlink and expose it to a quasi-democratic process of liking and plussing and hearting: Adding several links to a piece of text is usually not allowed. Hyperlinks are objectivized, isolated, stripped of their powers.

At the same time, these social networks tend to treat native text and pictures – things that are directly posted to them – with a lot more respect than those that reside on outside web pages. . . . A link to the pictures somewhere outside Facebook . . . are much less visible to Facebook itself, and therefore get far fewer likes. The cycle reinforces itself . . . Instagram – owned by Facebook – doesn’t allow its audiences to leave whatsoever. You can put up a web address alongside your photos, but it won’t go anywhere. Lots of people start their daily online routine in these cul de sacs of social media, and their journeys end there.⁹

There are billions of people on the Internet, but a few services capture or shape most of their activities. Take Facebook: it has 1.5 billion users, a billion of whom log in daily to see updates and news from the hundreds of people they have “friended” on the platform.¹⁰ Or consider Google: more than one billion people use the site to run more than three billion Google searches per day. Facebook recently announced a program encouraging publishers to upload articles to Facebook’s servers to make them appear faster to the end-users. Google is planning a similar gambit with “instant” articles of its own. As smartphones continue to claim an increasingly large share of Internet users, Google is also designing a new way to display pages on mobile devices.¹¹ Google’s new scheme would shift more power to the company; though, as with all the other transitions, it would offer benefits to users as well, which often serve to mask, or at least make palatable, the expansion of power.

For an increasing number of people, Facebook and Google are *the* Internet, or at

Zeynep
Tufekci

least the framework that shapes their experience of it.¹² These platforms own the most valuable troves of user data; control the user experience; have the power to decide winners and losers, through small changes to their policies and algorithms, in a variety of categories, including news, products, and books; and use their vast earnings to buy up potential competitors.

I talked with Derakhshan (online, since he is still in Iran) about his experiences, sharing my own research about the shift to a world of algorithmic walled gardens. Both of us are aware that current social media platforms reach many more people than the Internet did in the heydays of blogging. That is not the problem. Neither is it the existence of more frivolous or mundane content online; cute cat and baby images are part of the package. The problem is the shift in the architecture of the Internet. In ways both dramatic and subtle, the shift has begun to create new profound and far-reaching problems. In Derakhshan's words, a link is not just a link; it is a relationship. The power of the Internet comes from our relationships on it. And these relationships are increasingly mediated by the platforms that collect data about us; make judgments about what is relevant, important, and visible; and seek to shape our experiences for commercial or political gain.

How did we get here? And how much power is now concentrated in these platforms? The answers to these questions are connected and offer hints of possible alternative futures.

Legal scholar Lawrence Lessig has famously listed four forces that shape "cyberspace": law, norms, markets, and code.¹³ He compared his model to the offline world where law, norms, markets, and architecture play a major role in shaping society. Lessig analogized computer code, which defines how online platforms work, to the role architecture plays offline. Take the layout of

a city, for example: When residential and office buildings are separate, and people live in far-flung suburbs, there are social, political, and cultural consequences. Low walkability may contribute to unhealthy lifestyles. Or political polarization may increase while people segregate by income levels and race.

Online, computer code offers a similar structuring power. For example, Facebook requires mutual consent to interact, while Twitter allows people to "follow" someone else without being followed back. On Facebook, friending someone requires acquiescence on both sides: the person making the request and the person accepting it. On Twitter, any public account can be followed with just a click, without having to formally ask for permission. These structures are formed through decisions made by the people who run, administer, and create the code for these platforms, and are implemented by in-house coders, resulting in different social and political environments for each service. Facebook tends to have smaller networks made up of friends, family, and acquaintances, while Twitter is better suited for fan/celebrity relationships in which the few can be followed by the many. Online platforms are shaped not only by the code that structures visibility and access, but by computation and data as well. This combination gives online platforms powers for which there are no simple analogies in the offline world.

The massive accumulation of user data has been written about extensively.¹⁴ There is an increasing amount of data about everyone. More and more social, political, and financial interactions are performed online. More and more people carry phones that connect to the Internet and log their location and activities. Everyday objects are increasingly acquiring sensors that collect information even about passersby. Some of these data are accessed by governments for political purposes; some are used by com-

panies and advertisers for marketing. Financial institutions mine data to check credit-worthiness. Occasionally, the data are leaked, hacked, or otherwise released for reasons that can range from crime to politics to mischief. Ordinary people have very little idea about who holds what kind of data about them, or how the data are used. The amount of accumulated data and the asymmetry of power between the people who are monitored and surveilled and the platforms in which the data are held and mobilized is a significant problem, confirmed by polls revealing the public's great uneasiness about surveillance.¹⁵

However, the involuntary accretion of massive amounts of data about people is only the tip of the iceberg. In a networked society, computation brings another dimension of asymmetric power. Through techniques that can be loosely collected under the heading "computational inference" – the application of statistical methods, modeling, and machine learning to vast troves of data to make predictions – those who have gathered these data can infer from them information that has never even been disclosed.¹⁶

In other words, aided by computation, big data can now answer questions that have never been asked about individuals who are the sources of the data:

The advent of big datasets that contain imprints of actual behavior and social network information – social interactions, conversations, friendship networks, history of reading and commenting on a variety of platforms – along with advances in computational techniques means that political campaigns (and indeed, advertisers, corporations and others with the access to these databases as well as technical resources) can model individual voter preferences and attributes at a high level of precision, and crucially, often without asking the voter a single direct question. Strikingly, the results of such models may match the quality of the answers that were only ex-

tractable via direct questions, and far exceed the scope of information that could be gathered about a voter via traditional methods.¹⁷

The computational inference generated by machine learning takes place during the process of sifting through many varieties of data, with the proviso that the data are deep and rich enough. Inferring political variables about a person does not require their participation in overtly political websites or conversations. For example, Facebook operates mainly through likes: a one-click operation that signals a user's approval of a page, update, or person. The collection of these likes can be used to model, with surprisingly high statistical reliability, a range of outcomes, including "sexual orientation, ethnicity, religious and political views, personality traits, intelligence, happiness, use of addictive substances, parental separation, age, and gender."¹⁸

This type of analytic power can go beyond many of the traditional categories used by demographers and advertisers to profile the public. By using only their social media imprints (again, not directly asking questions of individuals), researchers have been able to identify people who are likely to become clinically depressed in the future, even *before* the onset of clinical symptoms.¹⁹ Much of this research is done with the best of intentions: for example, as early intervention for new mothers at risk for postpartum depression.²⁰ However, it is easy to see the downsides of making inferences using data in this fashion. Advertisers, for example, discovered that when women feel "lonely, fat, and depressed" they are more likely to purchase makeup, and that such women are ideal targets for "beauty interventions."²¹ In other words, women who are depressed and lonely can be more easily sold makeup. It does not take much imagination to see that advertisers will therefore want to use data gathered by online platforms to find out exactly who is feeling "lonely, fat, and depressed" and mar-

ket to these targeted women at exactly these times.

The increasing use of opaque computational methods known as machine learning – or “neural networks” – adds another layer of complexity to predictions made with big data sets. These are systems that “learn to learn” how to classify individuals (or whatever type of cases they are presented with) into various categories. Machine-learning systems are often provided with a “training set”: a database in which cases are marked with the correct answers.

For example, to train a machine-learning system, an employer might provide it with a list of employees he has classified as either “high-performance” or “unsatisfactory,” accompanied by social-media data about all employees in the database. Without receiving direct instruction or a recipe about what makes a worker either high-performance or unsatisfactory, the system learns the set of associations that are linked to each outcome, and how to use that knowledge to classify new employees. On the surface, this looks a lot like many other methods that employers use to discriminate among potential hires. But there is a twist: a machine-learning system often does not provide any human-understandable clues to why it classifies the way it does. In fact, if we knew exactly what it was doing, there would be no need for the “machine-learning” part: we could just program the criteria ourselves. In reality, though, all that a manager might know is that the system places potential hires into one category or the other, without any understanding of what parts of the social media big data set were used as signals for a particular outcome.

For all a hiring manager knows, such a system might classify applicants based on criteria such as statistical likelihood of experiencing depression in the future (even if undiagnosed at the time of evaluation) or the possibility of impending parenthood. It is well known that current hiring systems

incorporate many biases. However, if we use social media data churned through computational methods for hiring, we may move from imperfect hiring systems that we know discriminate against women, for example, to ones whose workings are hidden from us, but nonetheless still discriminate. This could mean using systems that discriminate *only* against women who are statistically likely to become pregnant soon. This type of discrimination would not be visible to employers because neither the women being hired nor the women not being hired would be pregnant at the time of the hiring, and because a machine-learning system does not display decision-making variables that are easily interpretable, even by its engineers. Social media platforms increasingly hold the kind of data that can be used in these ways.

While this combination of big data and computation obviously creates significant challenges, there are additional, equally daunting issues. When combined with the power of “code” as architecture, in the sense first identified by Lessig,²² platforms can also nudge behavior, quietly and imperceptibly, and sometimes in ways that are not directly visible even to the people who run the platforms. Facebook, for instance, uses an algorithm to order the news feed that shows its 1.5 billion users’ status updates. These may range from updates that are purely personal in nature to news articles. Increasingly, for many population segments ranging from younger people in developed countries to populations just coming online in poorer countries, Facebook has become the number one source of news.²³ In poorer countries, many people are not even aware that there is an Internet outside of Facebook,²⁴ and many others choose to stay completely within Facebook’s realm.²⁵ As David Clark explains in his essay in this issue, Facebook has helped ensure this through promotion of its stripped-down Facebook

app – o.facebook.com – which, in agreement with mobile service providers in many developing countries, does not incur data charges for users.²⁶

In my research, I have encountered many people whose Internet routine resembles the following: If on a desktop computer, a user launches a browser and types “Facebook” into Google’s search box, likely unaware that the URL bar at the top of the browser is a separate and faster way to get there. Google brings up Facebook as the first link, and the user clicks on Facebook and proceeds to interact mostly within the site. If using a mobile platform, which is increasingly the norm, a user will simply launch the Facebook app and rarely encounter the open Web at all.

This tendency to stay within Facebook is what gives Facebook’s architectural decisions such power, and invisibly so. In one study, 62.5 percent of users had no idea that the algorithm controlling their feed existed, let alone how it worked.²⁷ This study used a small sample in the United States, where the subjects were likely more educated about the Internet than many other populations globally, creating a potentially unrepresentatively low estimate. The news feed is a world with its own laws of physics, and the deities that rule it are Facebook programmers. In this world, some types of information are nudged and helped to spread more, while others are discouraged. There is great power in what we do (and do not) see from our friends and acquaintances, and increasingly, this is greatly influenced by platform design and code.

Facebook’s own research has shown the power of its designers’ architectural choices. In one Facebook experiment, randomly selected users received a neutral message to “go vote,” while others, also randomly selected, saw a slightly more social version of the encouragement, noting also which of their friends voted using small thumbnails of their profile photos. Matched with voter

rolls, that single message caused 340,000 additional people to turn out to vote in the 2010 U.S. congressional elections.²⁸ In another experiment, Facebook randomly selected whether users saw posts with slightly more upbeat words versus more downbeat ones: the result was correspondingly slightly more upbeat or downbeat posts by those same users. Dubbed the “emotional contagion” study, this incident sparked international interest in Facebook’s power to shape the user’s experience.²⁹

The power to shape experience (or perhaps elections) is not limited to Facebook; there are other powerful platforms. For example, Google rankings are hugely consequential. A politician can be greatly helped or greatly hurt if Google chooses to highlight or hide, say, a link to a corruption scandal on the first page of its results. A recent study showed that slight changes to search rankings can shift the voting preferences of undecided voters, and that these shifts can be hidden so that people show no awareness of the manipulation.³⁰

For a small taste of how platform choices affect the civic sphere, consider the case of the protests in Ferguson, Missouri, in August 2014. What started as a community shaken over the police killing of a young man under murky circumstances grew into major protests after the police responded to initial small-scale – and completely non-violent, according to journalists on the scene – protests by residents with overwhelming force, including the use of attack dogs and tear gas. A few national journalists, as well as ordinary citizens with smartphones, started tweeting from the scene of the initial protests. The burgeoning unrest and conflict soon grew into major Twitter discussions that later sparked the attention of the mainstream news media. About three million tweets were sent before the mass media began covering events in Ferguson. The nationwide movement that grew from these events is often referred to as the “Black Lives

Matter” movement, named after the Twitter hashtag.

However, on the first night of the protests, the topic was mostly invisible on Facebook’s algorithmically controlled news feed.³¹ Instead, the “ice bucket challenge,” in which people poured cold buckets of water over their heads and, in some cases, donated to an ALS charity, dominated the Facebook news feed. This was not a situation that reflected Facebook users’ lack of interest in the Ferguson protests; rather, it was an indication that it is hard to “like” – Facebook’s dominant algorithmic signal – such disturbing news, while it is easy to give a thumbs-up to a charity drive. Once a topic is buried by an algorithm, this becomes a self-feeding cycle: fewer people are able to see it in the first place, with fewer still choosing to share it further, causing the algorithm to bury it deeper. On Twitter’s platform, in which users see all posts from the people they follow in chronological order, the topic grew to dominate discussion, trending locally, nationally, and globally, catching attention of journalists and broader publics. On Facebook, it barely surfaced. Given the importance of online platforms and public attention to political movements, burying such news is highly consequential.³² Had our media been exclusively controlled by an algorithm in which “liking” were the main emotive input, the long and hard national conversation about race and policing in America that was generated by the Ferguson protests might have never transformed into a national movement.³³

How did we get here? Was it inevitable? Tracing this path requires combining and probing the two questions posed by Hani and Derakhshan, two people who blogged under repressive regimes and who were released from prison five years apart. Why is everyone on Facebook now? And why are these platforms killing the hyperlink as an

independent relationship between people? Why are they dictating who sees what?

Some aspects of the answer are deceptively simple, and at the same time deeply structural. The open Internet that held so much generative power took a turn toward ad-financed platforms, while the dangers lurking for ordinary users from the Internet’s open and trusting design were not counteracted, causing people to flee to safer and more user-friendly platforms. In combination, these two developments encouraged, enabled, and forced the creation of massive, quasi-monopolistic platforms, while incentivizing the platforms to use their massive troves of data with the power of computational inference to become better spy machines, geared toward ad delivery, the source of their financing.

From Wikipedia to question-and-answer sites to countless numbers of sites and blogs that provide a public service (but not payment for their creators), the Internet offers direct proof that people enjoy sharing their creative and personal output with others.³⁴ If there were ever a need to expand our conception of humanity beyond the restricted “homo economicus” who works only for his or her benefit, the explosion of user-generated content on the Internet has provided major evidence.³⁵ However, creative and altruistic output alone does not provide financing for servers, coders, and database management. As the public Internet scaled up and grew in numbers of participants, many websites faced a dilemma: whether to charge their users, or to sell users’ eyeballs to advertisers.

It was a crucial turning point: were people going to be the customers, or were they going to be the product sold? Almost all of the major platforms went with advertising. As Ethan Zuckerman, then a staff member of one of the Internet’s earliest user-generated platforms, tripod.com, explains:

Advertising became the default business model on the web, “the entire economic founda-

tion of our industry,” because it was the easiest model for a web startup to implement, and the easiest to market to investors. Web startups could contract their revenue growth to an ad network and focus on building an audience. If revenues were insufficient to cover the costs of providing the content or service, it didn’t matter – what mattered was audience growth, as a site with tens of millions of loyal users would surely find a way to generate revenue.³⁶

These decisions were made partly out of idealism: a free website-hosting platform like Tripod also allowed Thai dissidents to circulate otherwise censored content without worrying about paying for the site. It made more sense at the time to have ads than to charge users. But once advertising became the way to make money, almost everything flowed from it, especially when combined with another key feature of online platforms: network effects.

Network effects, also called network externalities, are the tendency of the value of some products or services to increase as more people use them, and to become less worthwhile when they are not used by others, even if the less popular product or service is objectively better, cheaper, faster, or more diverse in its offerings. For many online applications, everyone wants to be where everyone else is. This dynamic allows many online platforms that manage to get ahead of their competition to completely dominate their niche:

The more people own fax machines, for example, the more useful each one becomes. That is also why there is a single standard for fax machines – would you switch to a brand new, faster fax machine standard if there was nobody else you could fax with your machine? Research shows that the presence of network externalities trumps product preference or quality; many people will chose a service that has more users compared to the one that is otherwise better for them. Such

platforms, such as Facebook, tend to quickly dominate their market and become near-monopolies. This is also why everyone lists their wares on Ebay, where all the buyers are, and advertises on Google, where all the eyeballs go. The fact that a lot of people already have Facebook accounts means that considerations of network externalities will result in existing people staying put, or new people joining in anyway, even if they have qualms about the privacy issues.³⁷

While network externalities made it possible for platforms to become very large, the ad-financing model meant that a mid-sized platform, even one with hundreds of millions of users, faced great challenges, since ads on the Internet are not worth much.³⁸ An ad-dependent platform can only survive if it serves enormous numbers of people. For example, Wall Street’s investors have soured on Twitter because it only has about three hundred million users. For most products, hundreds of millions of users would appear to be a huge success. In an ad-financed online world, that’s barely enough to get by.

But there is one key path for online ads to become more valuable for platforms. If platforms accumulate a great amount of data on their users, and harness computational inference to “understand” them on behalf of their advertisers, then the ads, which have a higher chance of leading to a purchase, are worth a lot more. These advertisers could include both corporate entities selling products and political campaigns marketing politicians. Platforms can also use their architectural power to create an environment that is more advertiser-friendly. Until quite recently, for example, Facebook allowed likes as the only signal (aside from making comments) that users could send about a page or status update. While Facebook recently expanded choices in a few countries to include a few more “one-click” options such as “like,” “love,” “haha,” “yay,” “wow,” “sad,” and “angry,” the

Zeynep
Tufekci

expanded list is still heavily geared toward positivity, with only two that are typically associated with negativity: *angry* and *sad*.

Overall, many of the issues identified in this article are a direct consequence of this combination: Internet platforms are financed by ads that demand great scale, and they are fueled by network effects that allow such scale through the emergence of monopolies. These quasi monopolies then have incentives to collect and process vast amounts of data on their users to make the ads more effective for the advertisers, while also controlling the experience of the users to keep the platform advertising-friendly, and to keep the user from leaving the platform.

The other major development over the past decade from the user side has been the lack of attention and resources to ensure that the open Web – the one in which the hyperlink and address bar, rather than a closed platform and its algorithmic and architectural choices, dominate navigation – remains a secure and navigable place for ordinary users.

Many of the early protocols that defined the Internet were developed for use by a trusting, small, and closed community of academic and military research staff. However, on the current scale of billions of people, the Internet’s insecurity, and the proliferation of malware, spam, and untrustworthy sites, has caused many to retreat to easier-to-use, relatively safe platforms. The ad-financing model means that almost all commercial websites have installed extensive ad-tracking software on their sites, which is not distinguishable, in effects or operation, from malware dedicated to spying. Navigating the ordinary, open Internet now seems treacherous and feels slow (since the sites are loaded with ads and tracking software).

In 2014, for example, a massive vulnerability was found in “OpenSSL,” one of the

protocols that underpins almost all Internet commerce. The bug “heartbleed” allowed an attacker to read parts of a computer’s memory that the program should not ordinarily have access to, and to learn crucial private information, including stored passwords. While it is almost too ridiculous to believe, the OpenSSL architecture, used by about two-thirds of all web pages, including almost all major banks, is maintained by a group of only a dozen people, all but one of whom are volunteers.³⁹ The crisis with OpenSSL was but one example of critical parts of the Internet’s infrastructure that provide security for ordinary users being tended by almost nobody. There is very little energy or resources dedicated to tending the commons of the Internet, and the resulting environment has made ordinary Web navigation increasingly difficult and user data increasingly insecure. For regular users, remaining within trusted walled gardens, like Facebook or Google’s new proposed Web architecture, is a reasonable choice. This is exactly the scenario warned against by scholars.

This shift toward the walled gardens is only increasing as the next billions come online: people with less technical literacy, less powerful devices, shakier Internet connections, and often mobile-only access. In developing nations, the walled gardens of huge online platforms have many draws. Network effects means that their expatriate relatives and friends are most likely to be on the biggest platforms. A controlled environment makes the Internet more navigable. Bigger platforms offer better translation and localization services, something volunteer sites have more difficulty providing. Google helps order the chaotic, seemingly endless, choices effectively, while Facebook offers a way to manage the flow of information from a user’s social networks, albeit algorithmically curated within an ad-delivery platform. And thus, the Internet giants continue to grow, and have become

the dominant landscape of the Internet for most people.

In his prescient book *The Future of the Internet – And How to Stop It*, Jonathan Zittrain warned about these problems, and predicted that unless addressed, they would lead to the collapse of the open, generative Internet in favor of closed systems.⁴⁰ Legal scholar Tim Wu looked into past information systems and pointed out that many became dominated by monopolies.⁴¹ As early as 2003, Deborah L. Spar – now president of Barnard College – predicted that insurgent technologies would pass from the “pirates” that use technologies to disrupt order to the hands of powerful commercial and governmental bodies who use it to consolidate power.⁴² The Internet, in some ways, seems set on this path, although we have not yet passed the point of no return.

Because ad-based financing quickly devolves into large-scale, monopolistic systems working on behalf of advertisers, to change directions, we first must change how we finance the Internet’s platforms, including financing potential challengers to currently dominant platforms. Alternative models of financing were developed in the early days of the Internet, but these were quashed, in part because they may have been too early for mass adoption, but also because banks and websites resisted their implementation. Second, the Internet’s commons needs tending, which will also require substantial resources and financing as well. A global system whose security depends so much on volunteer work will, inevitably, become a difficult-to-navigate, insecure, and unpleasant experience, and will result in people retreating to safer platforms that cushion the user experience while also controlling it. Third, we must recognize that due to network effects, unregulated markets (one of the mechanisms of Lessig’s original four forces) do not work well on the Internet for certain kinds of platforms, including

many of the current tech giants. The influence of network effects is especially powerful for user-generated platforms, since what partially powers them is user investment. People have spent a lot of time and effort building up their positive feedback on eBay and cultivating their social networks on Facebook. It is unlikely that competition alone – even competition offered with better terms and services – can dislodge these powerful platforms, given the costs sunk into them by their users.

The path toward change is uphill, but the first step requires the public recognition of what dissidents in repressive regimes – often the canaries in mines – have already discovered: the power of the Internet derives from our ability to freely connect with each other. These developments are not changing only from one type of program or site to another; they are shifting to a new regime in which our relationships are mediated by forces trying to mine our data, mostly in order to sell a few more ads slightly more effectively, but also open to a host of other political uses.⁴³ From politics to culture, much power resides with owners of data, especially those possessing command of computation and online architecture.

It is not too late to change this path, but to do so requires an open-eyed and realistic look at the forces that have brought us here – financing models, the need for tending the security of the Internet’s commons, demand for usability, and the shift to mobile – and asking how to generate an alternative model that can scale-up. That demand still exists: the first billion Internet users have experienced, and remember, the admittedly chaotic early Internet, built upon the energy and euphoria of people discovering both information and each other. Now that the Internet is approaching three billion users, the question facing us is whether their Internet experience will much differ from a tightly regulated coffeehouse within a gigantic shopping mall.

Zeynep
Tufekci

ENDNOTES

- 1 He did not ask me to keep his identity secret, but I am not using his name on principle, to avoid my arguments getting tangled with his views as a result of Google searches run by clumsy repressive regimes.
- 2 Jennifer Preston, "Movement Began With Outrage and a Facebook Page That Gave It an Outlet," *The New York Times*, February 5, 2011, <http://www.nytimes.com/2011/02/06/world/middleeast/o6face.html>.
- 3 Wael Ghonim, *Revolution 2.0: The Power of the People is Greater Than the People in Power – A Memoir* (Boston: Houghton Mifflin Harcourt, 2012); and the author's private conversation with Wael Ghonim (2015).
- 4 Zeynep Tufekci and Christopher Wilson, "Social Media and the Decision to Participate in Political Protest: Observations From Tahrir Square," *Journal of Communication* 62 (2) (2012): 363–379, <http://doi.org/10.1111/j.1460-2466.2012.01629.x>.
- 5 Yochai Benkler, *The Wealth of Networks: How Social Production Transforms Markets and Freedom* (New Haven, Conn.: Yale University Press, 2007).
- 6 Steven Johnson, "Can Anything Take Down the Facebook Juggernaut?" *Wired*, May 16, 2012, http://www.wired.com/2012/05/mf_facebook/.
- 7 Rebecca MacKinnon, *Consent of the Networked: The Worldwide Struggle for Internet Freedom* (New York: Basic Books, 2012).
- 8 Hossein Derakhshan, "The Web We Have to Save: The Rich, Diverse, Free Web that I Loved – and Spent Years in an Iranian Jail for – is Dying. Why is Nobody Stopping It?" July 2014, <https://medium.com/matter/the-web-we-have-to-save-2eb1fe15a426>.
- 9 Ibid.
- 10 Don Clark and Robert McMillan, "Facebook, Amazon and Other Tech Giants Tighten Grip on Internet Economy," *The Wall Street Journal*, November 5, 2015, <http://www.wsj.com/articles/giants-tighten-grip-on-internet-economy-1446771732>.
- 11 Joshua Benton, "Get AMP'd: Here's What Publishers Need to Know about Google's New Plan to Speed Up Your Website," *Nieman Lab*, October 7, 2015, <http://www.niemanlab.org/2015/10/get-ampd-heres-what-publishers-need-to-know-about-googles-new-plan-to-speed-up-your-website/>.
- 12 Leo Mirani, "Millions of Facebook Users have No Idea They're Using the Internet," *Quartz*, February 9, 2015, <http://qz.com/333313/millions-of-facebook-users-have-no-idea-theyre-using-the-internet/>.
- 13 Lawrence Lessig, *Code: And Other Laws of Cyberspace, Version 2.0* (New York: Basic Books, 2006).
- 14 See, for example, Viktor Mayer-Schönberger and Kenneth Cukier, *Big Data: A Revolution that Will Transform How We Live, Work, and Think* (Boston: Houghton Mifflin Harcourt, 2013).
- 15 George Gao, "What Americans Think about NSA Surveillance, National Security and Privacy," Pew Research Center, May 29, 2015, <http://www.pewresearch.org/fact-tank/2015/05/29/what-americans-think-about-nsa-surveillance-national-security-and-privacy/>.
- 16 Zeynep Tufekci, "Engineering the Public: Big Data, Surveillance and Computational Politics," *First Monday* 19 (7) (2014), <http://dx.doi.org/10.5210/fm.v19i7.4901>.
- 17 Ibid.
- 18 Michal Kosinski, David Stillwell, and Thore Graepel, "Private Traits and Attributes are Predictable from Digital Records of Human Behavior," *Proceedings of the National Academy of Sciences* 110 (15) (2013): 5802–5805, <http://doi.org/10.1073/pnas.1218772110>.
- 19 Munmun De Choudhury, Michael Gamon, Scott Counts, and Eric Horvitz, "Predicting Depression via Social Media," in *Proceedings of the Seventh International AAAI Conference on Weblogs and*

Social Media (Palo Alto, Calif. : Association for the Advancement of Artificial Intelligence, 2013), Zeynep Tufekci
<http://www.aaai.org/ocs/index.php/ICWSM/ICWSM13/paper/viewFile/6124/6351>.

- ²⁰ Munmun De Choudhury, Scott Counts, Eric Horvitz, and Aaron Hoff, “Characterizing and Predicting Postpartum Depression from Shared Facebook Data,” in *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing* (New York : Association for Computing Machinery, 2014), 626 – 638, <http://doi.org/10.1145/2531602.2531675>.
- ²¹ Lucia Moses, “Marketers Should Take Note of When Women Feel Least Attractive,” *AdWeek*, October 2, 2013, <http://www.adweek.com/news/advertising-branding/marketers-should-take-note-when-women-feel-least-attractive-152753>.
- ²² Lessig, *Code: And Other Laws of Cyberspace, Version 2.0*.
- ²³ Amy Mitchell, Jeffrey Gottfried, and Katerina Eva Matsa, “Millennials and Political News,” Pew Research Center, June 1, 2015, <http://www.journalism.org/2015/06/01/millennials-political-news/>.
- ²⁴ Mirani, “Millions of Facebook Users have No Idea They’re Using the Internet.”
- ²⁵ World Wide Web Foundation, *Women’s Rights Online: Translating Access into Empowerment* (Washington, D.C.: World Wide Web Foundation, 2015), http://webfoundation.org/wp-content/uploads/2015/10/WomensRightsOnlineWF_Oct2015.pdf.
- ²⁶ See David D. Clark, “The Contingent Internet,” *Dædalus* 145 (1) (Winter 2016), 9 – 17.
- ²⁷ Motahhare Eslami, Aimee Rickman, Kristen Vaccaro, Amirhossein Aleyasen, Andy Vuong, Karrie Karahalios, Kevin Hamilton, and Christian Sandvig, “‘I Always Assumed That I Wasn’t Really That Close to [Her]’: Reasoning about Invisible Algorithms in the News Feed,” in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (New York : Association for Computing Machinery, 2015), 153 – 162, http://www.researchgate.net/profile/Kevin_Hamilton/publication/275353888__I_always_assumed_that_I_wasn't_really_that_close_to_her___Reasoning_about_Invisible_Algorithms_in_News_Feeds/links/553aa2fd0cf245bdd764475f.pdf.
- ²⁸ Robert M. Bond, Christopher J. Fariss, Jason J. Jones, Adam D. I. Kramer, Cameron Marlow, Jaime E. Settle, and James H. Fowler, “A 61-Million-Person Experiment in Social Influence and Political Mobilization,” *Nature* 489 (7415) (2012): 295 – 298, <http://doi.org/10.1038/nature11421>; and Jonathan Zittrain, “Facebook Could Decide an Election Without Anyone Ever Finding Out,” *The New Republic*, June 1, 2014, <https://newrepublic.com/article/117878/information-fiduciary-solution-facebook-digital-gerrymandering>.
- ²⁹ Lorenzo Coviello, Yunkyu Sohn, Adam D. I. Kramer, Cameron Marlow, Massimo Franceschetti, Nicholas A. Christakis, and James H. Fowler, “Detecting Emotional Contagion in Massive Social Networks,” *PLoS ONE* 9 (3) (2014): e90315, <http://doi.org/10.1371/journal.pone.0090315>.
- ³⁰ Robert Epstein and Ronald E. Robertson, “The Search Engine Manipulation Effect (SEME) and Its Possible Impact on the Outcomes of Elections,” *Proceedings of the National Academy of Sciences* 112 (33) (2015): E4512 – E4521, <http://doi.org/10.1073/pnas.1419828112>.
- ³¹ Zeynep Tufekci, “The Medium and the Movement: Digital Tools, Social Movement Politics, and the End of the Free Rider Problem,” *Policy & Internet* 6 (2) (2014): 202 – 208, <http://doi.org/10.1002/1944-2866.POI362>.
- ³² Zeynep Tufekci, “Algorithmic Harms beyond Facebook and Google: Emergent Challenges of Computational Agency,” *Colorado Technology Law Journal* [formerly *Journal on Telecommunications and High Technology Law*] 13 (2) (2015): 203 – 218; and Zeynep Tufekci and Deen Freelon, “Introduction to the Special Issue on New Media and Social Unrest,” *American Behavioural Scientist* 57 (7) (2013): 843 – 847.
- ³³ Tufekci, “Algorithmic Harms beyond Facebook and Google.”
- ³⁴ Benkler, *The Wealth of Networks*.
- ³⁵ Yochai Benkler, *The Penguin and the Leviathan: How Cooperation Triumphs over Self-Interest* (New York: Crown Publishing Group, 2011).

- As the
Pirates
Become
CEOs:
The Closing
of the Open
Internet
- ³⁶ Ethan Zuckerman, “The Internet’s Original Sin,” *The Atlantic*, August 14, 2014, <http://www.theatlantic.com/technology/archive/2014/08/advertising-is-the-internets-original-sin/376041/>.
- ³⁷ Zeynep Tufekci, “Facebook, Network Externalities, Regulation,” *Technosociology*, May 26, 2010, <http://technosociology.org/?p=137>.
- ³⁸ Zeynep Tufekci, “Mark Zuckerberg, Let Me Pay for Facebook,” *The New York Times*, June 4, 2015, <http://www.nytimes.com/2015/06/04/opinion/zeynep-tufekci-mark-zuckerberg-let-me-pay-for-facebook.html>.
- ³⁹ Dan Goodin, “Critical Crypto Bug in OpenSSL Opens Two-Thirds of the Web to Eavesdropping,” *Ars Technica*, April 7, 2014, <http://arstechnica.com/security/2014/04/critical-crypto-bug-in-openssl-opens-two-thirds-of-the-web-to-eavesdropping/>; and Jose Pagliery, “Your Internet Security Relies on a Few Volunteers,” *CNN Money*, April 18, 2014, <http://money.cnn.com/2014/04/18/technology/security/heartbleed-volunteers/index.html>.
- ⁴⁰ See Jonathan Zittrain, *The Future of the Internet – And How to Stop It* (New Haven, Conn.: Yale University Press, 2008).
- ⁴¹ Tim Wu, *The Master Switch: The Rise and Fall of Information Empires* (New York: Knopf Doubleday Publishing Group, 2010).
- ⁴² Debora L. Spar, *Ruling the Waves: From the Compass to the Internet, a History of Business and Politics along the Technological Frontier* (New York: Mariner Books, 2003).
- ⁴³ Tufekci, “Engineering the Public”; and Zittrain, “Facebook Could Decide an Election Without Anyone Ever Finding Out.”