The Root of the Matter: Data and Duty

Rules for the New Digital Economy Should Look to Old Common Law Traditions

November 2018

Author

Tom Wheeler,

Senior Research Fellow

Shorenstein Center on Media, Politics and Public Policy and Mossavar-Rahmani Center for Business and Government at the Harvard Kennedy School

31st Chairman of the Federal Communications Commission (FCC), 2013-2017

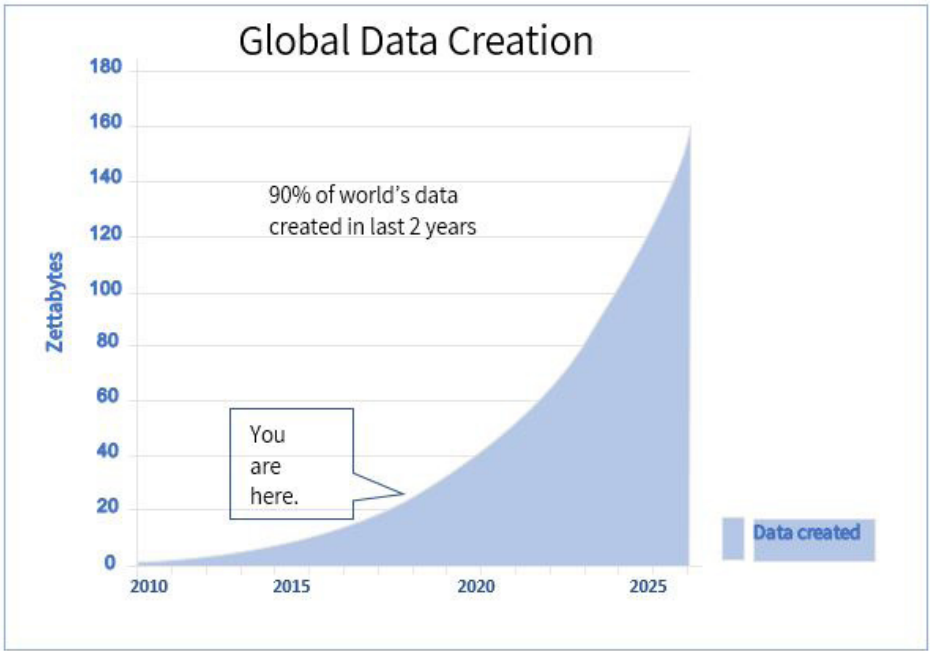
HARVARD Kennedy School
SHORENSTEIN CENTER
on Media, Politics and Public Policy

HARVARD Kennedy School
MOSSAVAR-RAHMANI CENTER
for Business and Government

There are 39 million books in the Library of Congress.[1] This impressive analog measurement pales in comparison, however, with the realities of the digital world. Every day connected computers create the data equivalent of three million Libraries of Congress![2] This startling fact defines the economics of our time. Serving as the foundation for many beneficial outcomes, the aggregation of this torrent of data also menaces the public wellbeing with threats as diverse as election interference and non-competitive markets.

It is important to recognize we are only at the beginning of the data aggregation era. Market intelligence firm International Data Corporation (IDC) forecasts that by 2025 the amount of data created and collected globally (both personal and machines talking to machines) will be over five times greater than today.[3] This startling forecast leads to two conclusions: (1) if we are concerned about the collection and manipulation of data today – get ready for even more, and (2) developing rules for handling this data and its impact on privacy and marketplace competition is a sooner-rather-than-later priority.

Much like the early days of the industrial era, when Rockefeller, Carnegie, Vanderbilt and other industry barons imposed their will, the companies driving today's digital economy are making their own rules. A century ago civic and political leaders stepped up to establish guardrails for industrial capitalism's management of hard assets, including protections for competitive markets, consumers and workers. The internet simply allows a new iteration of capitalism built on the soft assets of data-driven algorithms. The time has come for a new set of guardrails for information capitalism that



protect citizens and promote marketplace competition.

The framework for such policies already exists and is embedded in the principles of common law. Companies have responsibilities: a "duty of care" to not cause harm, and a "duty to deal" to prevent monopoly bottlenecks.

The time has come for a new set of guardrails for information capitalism that protect citizens and promote marketplace competition.

The harvesting of personal information – often without the individual's knowledge – infringes on the sovereignty of the individual and their personal privacy. Just as government once established rules to protect the collective good by assuring pure food and drugs, and clean air and water, we now have a collective interest in overseeing how the internet allows companies to collect and exploit personal information. Internet companies – both service platforms and the networks that deliver them – should have a "duty of care" as to the effects of their actions on personal privacy.

The subsequent use of that personal information has been cartelized to create new market-dominating anti-competitive forces. The data economy is no different from earlier economies where human nature and economic instinct created market-controlling bottlenecks. The dominant digital companies have a "duty to deal" so as not to block the competitive functioning of the marketplace.

We cannot overlook the many remarkable developments made possible by the companies of the digital economy. At the same time, however, it is time to reassert old truths and reestablish traditional duties to protect citizens and the competitive market. Such rules can ultimately benefit the companies and internet capitalism in the same way that past public policy decisions permitted industrial capitalism to flourish.

Harvesting Your Personal Information

> "Some of the most prominent and successful companies [in Silicon Valley] have built their businesses by lulling their customers into complacency about their personal information."[4]
>
> - *Tim Cook, CEO, Apple*

> "There is what I call the creepy line. The Google policy on a lot of things is to get right up to the creepy line."[5]
>
> - *Eric Schmidt, then CEO of Google*

Digital information is the most important capital asset of the 21st century. Twentieth century industrial assets were hard assets: products that were made in factories and sold in stores. Today's economy runs on the soft assets of computer algorithms that crunch vast amounts of data to produce as their product a new piece of information.

Digital information can make businesses more productive, such as helping jet engines function better or delivering produce to the grocery more efficiently. Data can also be threatening, such as the information gathered about the personal activities of each of us: our likes and dislikes; our habits and movements; our preferences and prejudices.

The collection and use of personal information have become so pervasive and promiscuous that oversight of these activities has become essential.

It is the aggregated personal information about each of us that allows countries such as Russia and Iran to attack our democracy. Coordinated misinformation campaigns begin with the precision of the information that online companies possess about each of us that is then used to guide influencers to susceptible targets.

It is the same aggregated personal information about each of us that also allows a handful of companies to attain dominant market positions at the expense of an open, fair and competitive marketplace. Because the currency of the new economy is data, whoever has the most can control important segments of the marketplace by denying it to others.

The collection and use of personal information have become so pervasive and promiscuous that oversight of these activities has become essential. "I'm not a pro-regulation kind of person," Tim Cook told a gathering, "But I think you have to recognize that when the free market doesn't produce the result that's great for society. You have to ask yourself: What do we need to do? And I think some level of government regulation is important to come out of that."[6]

At the heart of the issue is the debate over who owns the information collected about individuals? The companies assert that their collection, aggregation and algorithmic manipulation results in their ownership of the data. But what then of the raw material inputs? If the oil buried beneath my land is my property, then why isn't the information buried deep in my life treated similarly?

The privacy rights of Americans have evolved on an *ad hoc* basis. Congress has granted statutory rights (*e.g.,* privacy of medical records). The Supreme Court has identified constitutional rights (*e.g.,* privacy rights of the accused). Regulation has applied authority delegated by Congress (*e.g.,* the Federal Communications Commission's rules for telephone network privacy). Privacy can also be protected by contract (*e.g.,* a non-disclosure agreement). Unlike the European Union, however, prevailing law in the United States does not generally recognize that an individual's information is their personal property.

We have emerged into an era where the technology to collect and aggregate personal information has sped past the law.

We have emerged into an era where the technology to collect and aggregate personal information has sped past the law. The combination of digital technology and a drive for dollars has exploited the vagueness of American law. The time is upon us to eliminate that vagueness with policies based on a common law concept: the "duty of care" to not cause harm through one's actions. The internet companies – both networks and the services that ride on them – should have a "duty of care" as to the effects of their actions on personal privacy.

Surveillance Capitalism

Some of our personal information has always been available to other people. The bank knew how you spent your money. The postman saw your political leanings in your mail. The Division of Motor Vehicles knew what you drove, where you lived, and how safely you drove. But that information was scattered and unconnected.

In 1978 the security that resulted from the scattering of such information began to disappear. That year IBM introduced "relational database" software that allowed the information in one database to be correlated with the data in another to suggest a conclusion.[7] Such analysis soon became a staple of corporate management to look for patterns in the aggregated actions of individual customers or suppliers.

For a while the high cost of computers and data storage acted to protect the expectation of personal information security. Yes, your bank now had the ability to quickly compare your spending habits with your mortgage, but such analyses were still costly and institutionally encapsulated.

When Moore's Law drove down the cost of computing to put it on every desktop and ultimately in every pocket, and the internet connected those devices with a common *lingua franca,* the *ad hoc* protections to information sharing vanished. Suddenly, rivers of data were being created, connected and compared. New services sprung up with new relational capabilities that allowed consumers to access databases to answer questions, connect with friends, and purchase products. Each one of those interactions created new and very valuable information.

The early iterations of internet platforms took advantage of these

capabilities in a simple exchange: you told Google what you were looking for, you told Facebook about yourself and your friends, or you told Amazon about a product you wanted. In return, these services would use that information ("Tom is looking for…" or "Tom's birthday is…") for targeting advertisements and other messages. It was the digital version of the demographic targeting long used by publishers and advertisers.

The increased reliance on mobile devices to access the internet opened up a whole new pathway to the collection of personal information.

This simple trade became more complex as the services placed small files ("cookies") on your device to communicate back to them. Cookies weren't new, they were originally used by websites to remember your preferences, thus allowing the site to respond quickly or bring up your last interaction. As the digital companies concluded that their best business model was to sell targeted advertising, they started planting cookies to track the user across the web in order to build a more detailed file on that particular person.

What started as a digital variation of traditional demographic targeting soon became a far more powerful business tactic driven by the ability to track anyone across large parts of the web, combine that data with other collected data, and infer an amazingly accurate portrait of him or her.

Then things got worse.

The increased reliance on mobile devices to access the internet opened up a whole new pathway to the collection of personal information. It became possible not only to know what a person was doing online, but also where they were doing it, where they'd been, and where they're heading. Tracking leapt out of the online world to follow the real-world behaviors of otherwise private individuals across multiple internet accessible devices.

Such tracking has only continued to grow. After testifying before Congress, Facebook CEO Mark Zuckerberg filed written replies to questions he had been asked at those hearings. A sample of what he reported included the following information Facebook records about you:[8]

- Information from other devices in your home (computers, phones, connected TVs) as well as information from your internet service provider or mobile operator,
- Information about nearby WiFi access points, beacons and cell towers and their signal strength to aid in locating the user,
- Information on purchases you make on non-Facebook sites,
- Contact information such as your address book and (for Android users) call log or SMS log history,
- Information on how you use your phone's camera, including the location at which a picture is taken,
- Information on the games, apps or accounts you use,
- Information about when others share or comment on a photo of them or send a message.

We have come a long way from the simple transaction that answered

a question or connected you with a friend. The issue going forward is whether components of internet capitalism have become nothing more than surveillance capitalism?

Protecting Personal Information

The Privacy Act of 1974 established rules regarding the federal government's collection and use of personally identifiable information. Having identified the importance of protecting private information, the Congress in subsequent years applied the concept to specific marketplace practices.

The Video Privacy Protection Act of 1988 responded to the search of Supreme Court nominee Robert Bork's video rental records by making the release of such data illegal. The Driver's Privacy Protection Act of 1994 protected the disclosure of personal driver's license information after such data had been used to track users of abortion clinics. The Health Insurance Portability and Accountability Act of 1996 protected the privacy of medical records. The Children's Online Privacy Protection Act of 1998 outlawed data mining of children under 13.[9]

Typically, congressional efforts to protect privacy have been associated with specific instances of abuse. Yet, those same concepts have not been applied to the abuses emanating from the most powerful and pervasive information collection network in the history of the planet, the internet. The intervention of government is now necessary because the companies that collect our personal information lack incentive to fully protect our privacy.

For over a decade we have been told that the companies that profited from our information were, indeed, concerned about our privacy. That representation was based on the four great myths of internet privacy:

- Myth #1: "You are in control of your data" – By one count, Facebook CEO Mark Zuckerberg delivered this message 45 times during his two congressional hearings.[10] We are "in control" as much as we might be against an extortionist's threat. This time, the message is "We're holding your service hostage until you pay us with your private information." Even when we are told we can be in control, we often aren't. The Associated Press reports that after users' exercise "control" to turn off the tracking function of their smartphone, Google continues to track and store the consumer's locations.[11]
- Myth #2: "Privacy policies protect you" – In Orwellian doublespeak, the "privacy policies" that are made to sound as though they protect privacy are actually about permission to violate your privacy. The defense that "nobody reads it anyway" falls apart in the density of every online service's unique terms and conditions. Researchers at Carnegie Mellon University found the median length of the privacy policy from the top websites was 2,514 words. At a standard reading

rate, it would take 76 eight-hour work days – almost four months – to read the privacy policies of the websites visited by the average American.[12] What's more, these so-called "protections" can be – and are – changed at will by the companies. Relying on so-called "privacy policies" for protection is like hiring the cat burglar to guard the jewels.

For over a decade we have been told that the companies that profited from our information were, indeed, concerned about our privacy.

- Myth #3: "The trade of value for service" – The economic equilibrium that once may have established the exchange of "free" services for targeted information no longer exists. What began as "give us relevant information in exchange for service," has become "we want all your information, including what you are doing when you're not interacting with us, and while you're at it, we want the information of your friends." The information collected is not just about your online behavior, but also your real-world behavior, including your location over time, the places you frequent, even the floor of the building you are on. It is an exchange that, as Senator Mark Warner (D-VA) has explained, has "less price transparency…than there is in health care. And look how screwed up that market is."[13]
- Myth #4: "The information is anonymous" – In the world of large data bases and powerful computers, there is no such thing as de-personalized data. Even when the data collected is "user anonymous," such anonymity disappears in the milliseconds it takes a computer algorithm to look at the multiple pieces of data collected about an individual and connect the dots. Researchers at MIT and the Université Catholique de Louvain in Belgium found that supposedly anonymous cellphone data could be associated with a specific individual 95 percent of the time using only four data points.[14] For comparison purposes, to identify an individual from a fingerprint requires identification of 12 different inputs from points on the print.[15] This ability to identify, or at least infer, the specific individual extends to their name, phone number, birthdate, and in many cases credit card number.[16]

Shrouded by these myths is the simple truth that consumers have lost control of that most personal of assets: their own information.

Permission, Not Prohibition

The rights of individuals regarding their personal information is a 21st century civil rights issue. Consumers know they have lost control of their personal information. A 2017 survey found 70 percent of Americans lacked confidence that their personal data is private and safe from distribution without their knowledge.[17] A 2018 poll discovered that 95 percent of Americans believed their personal privacy "now rivals our Constitution's most basic rights" such as

the First Amendment.[18]

Previously, such an awareness about the vulnerability of personal information, whether video rental records or DMV information, triggered new privacy protections. We are once again at such an inflection point.

Returning control of their information to consumers has been embraced by both the right and left of the political spectrum. Alt-right political strategist Steve Bannon describes it as the recovery of each citizen's "digital sovereignty." He cites it as one of the three principles that will drive a forthcoming populist rage.[19] The political left has adopted a similar theme, calling it "digital feudalism." In Medieval times, feudal lords confiscated the output of their serfs – an individual's labor; today, digital lords confiscate the output of individuals – their information.

The rights of individuals regarding their personal information is a 21st century civil rights issue.

There can be no doubt about the wondrous new capabilities that have been made possible using digital information. From the ordering of a pizza, to medical research without beakers and lab rats, we are significantly better off because of the data-based innovations the internet has made possible. While applauding the success delivered by the innovative use of personal data, we cannot be blind to how the information economy has created new and unique problems that appear to be durable in their nature.

In considering the rights of individuals and their personal information, it is important not to be trapped into false debates. This is not a zero-sum choice between unfettered access and usage or the disappearance of online services. Developing policies about the use of personal digital information should be a matter of permission, not prohibition.

In 2016 the Federal Communications Commission (FCC) took up the matter of privacy for the personal information transiting digital networks. It was familiar territory for an agency that had for years enforced privacy on the telephone network. The telephone rules had been simple: without the consent of the consumer, the companies were prohibited from sharing information about a telephone call, including the information that set up the connection.

The privacy of a network connection is especially important since the network company sees all of a consumer's traffic – where they are going and at least part of what they're doing. The telephone rules provide, for instance, that if a consumer uses their smartphone to place a phone call to Air France, the phone company cannot sell that information to a French tour or hotel company. The 2016 FCC rule simply extended the telephone privacy concepts into the digital era so that if the same person uses the same smartphone to access the Air France web site, the network company could not sell that information to a French hotelier or tour operator.

The FCC's protections didn't last long, however. Following the election of Donald Trump, at the urging of both the networks and the platform companies, the Republican-led congress repealed the FCC's rule. What's more, the congress wrote into law a prohibition that the FCC could never again enact

similar rules.

When the Trump Administration and Congress repealed the FCC's privacy protections, the state of California stepped in. In 2018 the state legislature filled the void created by the digital companies' demand for no regulation. California consumers, at least, would have their own privacy protections. Immediately the networks and platform companies rushed to congress in search of a better deal that would preempt the state's action. The CEO of AT&T called on congress to "step up" and create privacy "rules of the road."[20] The fact that his company had recently worked to repeal national rules for networks that could have been the model for rules for all others seemed to be an undiscovered irony.

The creation of online privacy protections for American consumers should be based on three building blocks: transparency, control and responsible forethought.

Nonetheless, the call for uniform national rules applicable to the entire digital economy reinvigorates the debate over the components of meaningful privacy protections. The creation of online privacy protections for American consumers should be based on three building blocks: transparency, control and responsible forethought.

Transparency: Consumers should know what is being collected, how it is being collected, and how the information is being used, including what information is being stored for subsequent reuse. The digital companies, seeing the writing on the wall, have begun to embrace transparency, albeit under their own definition

But simple transparency is not a pure-play solution to the privacy challenge. The companies may now embrace what they call "transparency," but telling someone what you are about to do to them is not justification for the act itself. This is especially true when the digital companies can unilaterally change their data collection and usage policies simply by providing notice of the change.

Informing consumers about the data being collected is not absolution. Specific disclosure as to how the data is used and to whom it is available are essential to the meaningful enrichment of any transparency. Most important, however, is giving the consumer control over their own information and establishing for the companies specific expectations about their activities and responsibilities.

Control: Ninety-two percent of Americans believe companies should have to get permission before sharing or selling their online data.[21] Today so-called "consent" can be coercive (*i.e.,* unless you agree you can't use the app) or buried in thousands of words of legalese. Being able to "opt-out" after data collection has begun is not adequate privacy protection, it merely shifts the burden to the consumer to attempt to recover their own privacy.

As noted previously, there is a legal debate whether information about an American citizen is that individual's property. If two people have a conversation, for instance, who "owns" that information? When an individual has a digital interaction with a company, is that "conversation" treated similarly?

At the very least, however, the consumer should have control over

whether the "conversation" takes place at all. This means consumers should have up-front opt-in control of what information is collected and how it is used.

Technically, the job of a microprocessor is to inexhaustibly collect and compute information. When coupled with the ubiquitous connectivity of the internet, the result is a tsunami of data that is created and collected whether it is needed or not. Because the cost of storing such data is virtually nil, and the potential to enhance that data with other data is remunerative, the temptation to collect more data than is needed for the operation of a product or service is intense. There need to be guardrails on the unlimited power of the technology to collect information and the unbridled incentive of companies to accumulate that information.

Consumer control of his or her information should also include the ability to move that information to a different platform or service. In the subsequent discussion of market control, we will explore such an activity's competitive effects. Part of the digital "conversation," however, should be the consumer's right to export information from one database to another.

The computer science that created the capability to invade personal privacy can also be harnessed to build tools that consumers could use to protect their privacy. The economic incentive to build such apps is absent, however, because it is the use of the information, not its protection that is remunerative.

It is possible, for example, to build a smartphone application that identifies all the information the device is collecting. Knowing this, the consumer can choose who should have access to that data, and on what terms. Some information, for example location data, is essential for the operation of a wireless network. Once such data exists, however, the phone company or the platform company that captured it is currently free to use it for other purposes. If that data's provenance is returned to the consumer, then the consumer can decide what those other purposes should be.

Digital Forethought: Mark Zuckerberg was candid in his congressional testimony when he said the design of digital platforms often proceeded without consideration of the effect of that design. "We didn't take a broad enough view of our responsibility," he told the United States Senate. [22]

What has been missing thus far in the internet era has been exactly that kind of planning ahead to identify the possible effects of a specific digital activity. Asking the question, "Do we understand the implications on personal privacy of what we are building?" should be a threshold question in the creation of digital services, not an afterthought.

When Facebook designed its Application Programming Interface (API) – the software that allows different programs to work together – its lack of forethought allowed companies such as Cambridge Analytica to access the personal information of 87 million consumers without their permission. To Facebook's credit, it subsequently improved its existing privacy design review process. The

fact that such consequences were not realized until after harm had been done is illustrative, however, of the importance of privacy being the default assumption in the application of digital technology.

Google promotes how, "Google products and features cannot launch until they are approved by the specialists in our Privacy and Data Protection Office."[23] It is a responsible step, but the issue remains as to what policies the office is following to make its determination.

Requirements for transparency, control and forethought aren't revolutionary. At their heart is the common law concept of the "duty of care."

Ann Cavoukian, Privacy Officer for Ontario, Canada, working with the Dutch Data Protection Authority, first surfaced the concept of "privacy by design" in the mid-1990s.[24] The European Union's General Data Protection Regulation (GDPR) incorporates the idea, also known as "data protection by design and default." It means that privacy protections should be embedded in digital designs from the outset as an essential component of what is being delivered. This also includes the responsibility to collect and store only the minimum amount of data that is necessary for the provision of the specific service.

Old Duties Made New

Requirements for transparency, control and forethought aren't revolutionary. At their heart is the common law concept of the "duty of care."

Underlying the legal principle of negligence is the expectation of reasonable care being exercised to anticipate and mitigate the potential harm an activity might impose. Such a duty to care should be applicable to the activities of digital companies as they collect and exploit private information.[1]

Over 150 years ago, the world's first high-speed network, the railroad, imposed a similar set of challenges on its time. A boon to commerce and communications, the steam locomotive nevertheless brought with it adverse consequences.

The railroads, as state-chartered entities, were able to assert the state's right of *eminent domain* to confiscate the property of land owners for track rights-of-way. As the iron horse crossed the private property it belched hot cinders from its stack, some of which landed on the barns and hayricks being passed and set them ablaze. New risks were also created for railroad workers, individual citizens and animals around the tracks. The new network may have asserted property rights, but – like the new information realities of the internet – it had also created new responsibilities to mitigate the problems it was generating.

The standard of negligence – a duty of care – became the test. Had the railroad taken reasonable steps to mitigate the threats it had created? Had, for instance, a screen been installed on the smokestack to keep the hot cinders

1 Harvard's Jonathan Zittrain and Yale's Jack Balkin have proposed making online platforms "information fiduciaries," with the responsibility to act in the consumer's best interest. It is a concept consistent with the duty of care.

from escaping?

The idea of a digital duty of care should similarly apply to the protection of individual data privacy. The topics that were just discussed define such a duty of care:

- Simplified information about the consumer data being collected,
- True consumer control over the collection of their information, including its sharing with a third party,
- True consumer control over the usage of their information, including the right to review and edit appropriate information,
- Ending coercive collection that ties use of an essential service, or a service for which there is no competitive choice, to mandatory access to consumers' information,
- Product design that anticipates its effect on the privacy of users and collects only the data necessary for the product to function,
- Protection of consumer data after it has been collected and stored,
- Consumer-activated portability of data stored in one database to another,

The activities of individual states such as California and international organizations such as the European Union have combined to incent the digital network and platform companies to overcome their previous antipathy towards privacy regulation and ask Congress to develop a standard national privacy policy. Of course, the companies want that policy to be "light touch" so as to allow them as much leeway as possible.

There already exists, however, a bedrock concept that should guide the development of such policies: the "duty of care." The roots of common law have established expectations as to our responsibilities to each other. New technology does not alter those responsibilities.

The Economics of Information and Market Control

"Google-Facebook Dominance Hurts Ad Tech Firms, Speeding Consolidation"
- New York Times headline, Aug. 12, 2018

"I wouldn't say the internet has failed with a capital F, but it has failed to deliver the positive, constructive society many of us had hoped for."
- Sir Tim Berners-Lee, father of the World Wide Web[25]

The use of personal information is the entry drug to digital addictions and marketplace dominance.

The use of personal information is the entry drug to digital addictions and marketplace dominance. Internet companies have become addicted to the use of personal data to drive software algorithms for the highly profitable targeting of individuals. Because the specificity of these algorithms is determined by the depth of the data that goes into them, the companies that collect personal information have every incentive to lock it up for their own exclusive use and then to use the aggregated information as a tool with which to expand their power in markets.

The internet, a decentralized collection of interconnecting networks, early-on promised to bring freedom and choice – instead it has created a new centralized power to inexpensively capture information in a manner that tends toward the creation of new bottlenecks to the competitive functioning of the market. Common law long ago developed a concept to counter such bottlenecks: a "duty to deal." When someone controls access to a fundamental asset, they have an obligation to make that asset openly available, not for free, but as an asset whose openness can benefit all.

Digital Economics

Digital technology has been brutal to the companies and workers of the 20th century. An astonishing 52 percent of the Fortune 500 companies at the end of the 20th century no longer exist.[26] The evolution of the world's most valuable companies from the turn of the century to today tells the basic story about the transition from industrial economics to information economics and the high profit margins made possible by digital technology.[27]

In the industrial economy of the 20th century, to produce an additional product or to stock it on the shelves required a substantial additional investment. When GE wanted to build one more locomotive, it had to buy the steel and hire the workers to fabricate it. Exxon had to drill new wells and build or expand refineries. Walmart had to buy the products to increase its inventory. Citibank had to pay a price to acquire more money to lend.

The harbinger of the digital future at the turn of the century was Microsoft. Selling an additional piece of software was virtually costless; there was nothing new to build or acquire, just go to the file in which the item is stored and send a copy over the internet to the customer. The work on the product had already been done, and the internet delivery was virtually costless. It was an

early demonstration of how digital products have almost zero marginal cost.

The internet platform companies have improved that model even further. The business model at the core of the consumer-facing internet is premised on the virtually costless capture of digital information about individuals, then monetizing it through the delivery of targeted messages. Being able to pay virtually nothing for a product that generates sizeable revenue has to be the world's best business model.

Being able to pay virtually nothing for a product that generates sizeable revenue has to be the world's best business model.

What's more, unlike many industrial age goods in which the value of an asset tended to decrease over time, digital information shows implications of becoming more valuable as it is combined with other data to create even greater identification specificity. An industrial asset such as steel, for instance, would be stamped out into a fender that would begin depreciating. An information asset, in contrast, can appreciate in value when an algorithm combines it with other pieces of data.

The world's most valuable companies have successfully taken advantage of this digital redesign of economic activity. Apple, originally a 20th century hardware company, is also evolving to a digital platform (iTunes, App Store, iCloud, etc.) and its market capitalization has responded. Alphabet, the parent of Google, is the world's largest digital advertising company and gathers information through the world's leading browser, the number one mobile operating system, and the most popular search engine. The early pioneer Microsoft has also re-engineered itself into a data collection engine, building the collection of user data into its Windows10, for instance. Amazon simply knows more about our buying habits than anyone else. And, of course, the company that rounds out the top five most valuable companies in the world – Facebook – catapulted from a dorm room to such heights through the low-cost aggregation and high-value reuse of personal information.

World's Highest Market Capitalization Companies 2001-2018

The valuation of these companies reflects not only their low marginal cost/high marginal profit, but also their ability to use information dominance to achieve marketplace dominance.

Market Dominance

The five most valuable companies in the world know, or profit by enabling others to know, specific information about each of us. Having aggregated this information, the data is hoarded to create a bottleneck that maximizes its value by controlling its usage.

Apple tells us it does not sell access to personal information. However, the apps that run on the Apple operating system do, and the revenue they generate is the company's fastest growing line of business.[28]

Digital companies transform this kind of aggregated personal information into market power by controlling access to that information.

Google knows what people are interested in. Over 60 percent of all U.S.-originated search queries come through Google,[29] and Google's Android operating system that is used by 85 percent of the world's smartphones feeds even more information.[30]

Microsoft has been reinventing itself to monetize the information it collects about its users. The ubiquitous Windows10 operating system collects user data, and the Bing search engine, while a distant second to Google, still generates almost a quarter of all search queries and accompanying data.[31]

Amazon knows what you buy and uses that information to suggest what you should buy next. The result is a market mega force with revenue greater than Google and Facebook combined.[32] Growing exponentially in online retail by harnessing information and scale economies, Amazon is creating the next generation of retail, integrating the information it controls across not just the web, but also social media and even brick and mortar physical outlets.

Facebook knows who you (and your friends) are. First, because you tell them as a condition of receiving service. Then Facebook adds to that data through ownership of four of the top five mobile apps – Facebook, Messenger, WhatsApp, and Instagram – and the information they generate.[33]

Digital companies transform this kind of aggregated personal information into market power by controlling access to that information. Google and Facebook, for instance, control 48 percent of local digital advertising.[34] This is not because they are a part of the local community, but because their ubiquitous collection allows them to know more about the members of the local community than even the neighbors. Knowing and controlling such information allows these aggregators to siphon advertising dollars from the local media.

The impetus to become such a bottleneck is as old as commerce itself. Since open markets are competitive markets, economic incentive pushes for profit maximization through competition-thwarting bottlenecks. What is different in the information era is the changed characteristic of the asset being exploited by the digital bottleneck.

The use of industrial era raw materials was typically a use-once-and-its-gone activity. The coal that powered the factory was of limited reuse value. The digital raw materials that power the information era tend to be a store-and-use-often asset that demonstrates an ability to appreciate in value as it is combined with other data to provide an ever-more granular description of reality.

In the offline world, demand is created by new design or new features. In the online world, the aggregation and control of data not only drives demand, but also becomes a winner-take-all tool that can diminish competition.

Like their industrial counterparts, digital companies enjoy the benefits of scope and scale economies on the supply side. What has changed is that they also enjoy control on the demand side of the marketplace through the exercise of what economists' call "network effects."

Network connectivity has always been a matter of network effects – how the value of a service increases as its users expand. The first telephone was useless until there was another telephone with which to connect. Each additional telephone connected to the network further increased the value of the network – the "network effect." The addition of an incremental user of a digital service has a similar value creation effect. In the digital world, however, network effects are magnified. Not only does each additional user improve the experience for other users, but it also increases the economic value of the platform because the data the incremental user generates further enriches the targeting capabilities of the data the company already has.

The demand for data is driven by the precision of the data. Since targeting precision increases with the amount of data available, those able to enjoy network effects to increase their targeting capability can increase the demand for their product while decreasing the demand for others that do not have such large databases and thus lack such precision.

In the offline world, demand is created by new design or new features. In the online world, the aggregation and control of data not only drives demand, but also becomes a winner-take-all tool that can diminish competition.

Facebook and MySpace, for instance, were competitors in the early days of social media. Facebook ended up winning an industrial age-style contest for a better user experience. That win allowed network effects to kick in that created a data domination machine that assured no other company could challenge Facebook.

The first flywheel in the dominance engine is how as more users came on Facebook they generated more information. This additional data, in turn, enriched the company's targeting ability. By using that targeting to give users what they like, those consumers stayed on longer, were exposed to more paid messages, and in doing so created more information, which improved the targeting. It was an ever-repeating virtuous cycle. If today a new Mark Zuckerberg wanted to challenge Facebook the way the original Mark Zuckerberg took on MySpace – *i.e.,* with a good old-fashioned better product – the effort would require equivalent targeting precision attainable only through equivalent network effects.

The second flywheel of the dominance machine is the cash stockpile it makes available to buy off the competition. If a new company starts to succeed with a new idea that could threaten the incumbent, the big companies either exploit their large databases to power a copycat service with greater precision, or use their immense economic muscle to simply acquire the startup and fold it into the juggernaut. Facebook even purchased a company called Onavo whose software tracks web traffic to identify new services that might threaten Facebook.

Controlling markets by controlling data could extend into controlling the future.

Using Onavo, for instance, Facebook was able to quantify how Instagram was a threat. As one tech journalist observed at the time, Facebook "knew that for the first time in its life it arguably had a competitor that could not only eat its lunch, but also destroy its future prospects."[35] Facebook paid $1 billion to purchase the two-year-old company, bringing it under the corporate tent before Instagram could grow into a serious challenger. Enriched with Facebook's data, Instagram has become an important source of Facebook's growth.[36]

The dominance engine rolls on as data repositories fuel the expansion of these companies into new heretofore competitive markets. Amazon, for instance, uses information about who buys a particular genre of book to inform production decisions for its Prime Video service, available to those who pay an annual fee for free shipping. Not only do Prime subscribers buy 3-4 times as much as non-Prime customers,[37] but also each transaction – whether for books or razor blades – further enriches data about the individual customer, pointing the way to other activities in which the data can be applied. When Amazon's data on each of us takes them into the grocery business, as it did with the $13.4 billion acquisition of Whole Foods, the power of data dominance to potentially control markets only grows.

Controlling markets by controlling data could extend into controlling the future. That the dominant platform companies are heavily engaged in the development of artificial intelligence (AI) comes as no surprise. Already AI applications such as predictive text, voice commands, and facial recognition expand the dominance of the companies. But when AI is recognized as nothing more than the algorithmic manipulation of huge amounts of data to reach a highly probable conclusion, then those who possess such huge amounts of data stand to dominate AI and the future it represents.

Today, digital companies aggregate data about past actions to influence future decisions. When that aggregated data becomes the input to artificial intelligence, the companies that dominate data have an open pathway to dominate the automation of our lives.

Digital dominance begins with vast amounts of unshared data. In the industrial era, rules were established to protect marketplace competition by preventing market dominance. How to protect and promote competition in the information era is our new challenge.

The Principle of Openness

The internet was designed to be an open network. Based on open protocols, a "network of networks" was created in which any computer or network could become part by simply following a common set of rules.

The nondiscriminatory openness that created the world's most important network thus falls prey to gatekeepers whose business plan relies on the ability to close their assets in order to discriminate.

The technical rules that enable the internet were themselves collectively developed in an open process. The so-called "multistakeholder process" gave everyone a seat at the table and managed for consensus. After creating the internet the open development process continued for the network's technical standards, but fell apart for the commercial operation of the networks and the platforms using the network. Those who benefitted from the network's technological openness created their own closed business models.

The name "internet" itself – a shortened version of the original term "internetworking" – connotes the essentiality of digital information being able to move freely. Prior to the standards of the internet, there were multiple incompatible digital networks, each using proprietary protocols to create a business operating within a company-controlled "walled garden." Internet Protocol (IP) tore down those walls by creating a common *lingua franca* for the operation of digital networks.

While the internet itself results from the strict adherence to operational rules, there are no similar rules for those who connect to the internet and take advantage of its openness. At one end, an internet that is open in its "middle mile" as it knits together diverse networks into a whole, becomes closed in the "last mile" to the consumer. At the other end, the platforms that feed into the internet close off both their information and the secret decision-making of their algorithms.

The nondiscriminatory openness that created the world's most important network thus falls prey to gatekeepers whose business plan relies on the ability to close their assets in order to discriminate. The networks and platform services that today connect to the internet have exploited its open distributed digital architecture to reconstruct the kind of walled gardens the technology of the internet was designed to abolish.

On the network side, the wired and wireless networks that deliver users the last-mile to and from the internet won a decade-long net neutrality battle when the Trump FCC abolished the agency's 2015 Open Internet Rules. The companies are now free to discriminate amongst the traffic that openly arrives from the internet.

The absence of net neutrality means the traffic that has travelled the open middle mile of the internet on a non-discriminatory basis can suddenly hit a wall where the last mile networks can discriminate to favor one piece of traffic over another. When a wireless network company, for instance, delivers video content from a service it owns for free, while charging consumers data

rates to use a competing service, it has closed the openness of the internet that delivered both data streams without discrimination.

At the other end of an internet connection, the openness of the middle mile and the *lingua franca* of IP allows platform services such as Google or Facebook to easily reach consumers and siphon their data. Then that data is locked up behind the closed protocols of those platforms. These companies lobbied strenuously for net neutrality rules to keep the last mile open to consumers, yet they close their own activities in order to practice their own form of discrimination.

Openness is the proven initiator of innovation. Arguably, the internet's very existence is the result of opening up assets – both patent and network assets – that were being hoarded by a dominant company.

AT&T, the dominant network for most of the last century, created Bell Laboratories to research new technologies. It became one was one of the world's great centers of innovation. The laser, cellular phones, the solar cell and other technological breakthroughs happened at Bell Labs. But because AT&T controlled the patents protecting these developments, many discoveries were kept locked away lest they conceivably impact the core network telephone business.

In 1930, for instance, a Bell Labs engineer named Claude Hickman developed magnetic tape and built the first telephone answering machine. AT&T ordered cancellation of the project because management feared that the ability to leave a message would decrease the number of telephone calls.[38]

In 1956 AT&T settled an antitrust suit filed by the U.S. Department of Justice. One of the terms of that consent decree was AT&T's agreement to license its patents. The approximately 8,600 pre-decree patents would be licensed for free, while those that followed would be open upon payment of a reasonable fee.[39]

Included among the patents were two technologies seminal to the development of the internet: the transistor and the modem. The transistor – a small sandwich of silicon and resistors – replaced the vacuum tubes being used for on/off binary switching in the era's giant computers. Had it not been for the compulsory licensing of the AT&T patents, the development of increasingly smaller and more powerful microprocessors – the heart of the internet – by companies such as Intel would have been much more problematic. Gordon Moore, a co-founder of Intel, described the requirement that AT&T open its patents as, "One of the most important developments for the commercial semiconductor industry," a decision that allowed the semiconductor industry "to really get started."[40]

The transmission of the internet's digital signals also was a result of openness policies. AT&T had developed modem technology that converted digital signals to analog so they could be transmitted over the telephone network

(the word "modem" is a contraction of **mo**dulate-**dem**odulate for its ability to turn digital pulses into analog signals and back again). While the consent decree may have opened access to the modem technology, AT&T refused to allow non-AT&T devices to connect to the network. A U.S. Court of Appeals decision overturned the FCC rule protecting this practice in the 1956 *Hush-a-Phone* decision. The openness of the patents thus joined with the openness of the network to create the technology and the environment that birthed the internet.

The concept of openness for fundamental economic assets is not revolutionary; we have seen it before in history – multiple times.

As the dominant companies of the digital era hoard their assets, the application of remedies similar to those imposed on AT&T in the mid-1950s is getting a new look. The opening of AT&T's patents encouraged innovation and competition; the opening of closed digital assets, it is argued, could have a similar effect. Such a mandate to open key digital assets in the public interest is a challenging undertaking. Security and privacy issues would need to be addressed as a duty to deal does not override the responsibility to secure an individual's privacy.

Yet, there are examples of government mandating the opening of proprietary databases with consumer consent. In the U.K., for instance, open banking regulations require the nine largest financial institutions to allow third-party access to their customer information databases through an open and secure application programming interface (API).[41] The consumer's digital information remains an asset of the financial institution, but access to the bank's database is made available to third parties so they may develop new services. A third-party app, for instance, can now aggregate an individual's multiple accounts (banking, checking, credit card, mortgage, etc.) into a single database to better understand options and rationalize costs without having to be satisfied with whatever the bank is offering. In another application of the new rule, a competitive service can access the database of financial institution to send and receive payments that previously could only be done through that institution. The initiative is overseen by the Open Banking Implementation Entity (OBIE) which the government required the financial institutions to create.

In the United States, long before the issue of dominant digital databases, the FCC told mobile phone companies they could not hoard the phone numbers assigned to their customers. Previously, the companies used their control of the phone number to discourage consumers from enduring the "tell all your friends" hassle of switching to another company and being forced to use another phone number. When the FCC forced the companies to open their databases to allow a number to be ported from one carrier to another, consumers were empowered and competition increased.

The concept of openness for fundamental economic assets is not revolutionary; we have seen it before in history – multiple times.

Old Duties Made New

As our ancestors came to grips with the new technologies of the mid-19th century, they embraced the common law precedents.

The establishment of common law principles helped 15th century England emerge from medieval feudalism. One of the basic concepts developed in the common law was to open access to bottlenecks through a "duty to deal." Fundamental activities – a hostelry or a river ferry, for instance – had the responsibility to accept all comers, not just those the proprietor chose to serve. The internet is the 21st century river ferry: a fundamental activity upon which depends the economic activities of others. Similarly, just as the ancient innkeeper controlled access to food that should not be denied others, the digital platform companies control access to the data sustenance of the digital economy.

Such common law responsibilities enabled the mercantile era to bloom in the 16th through 18th centuries. Even the era's leading free market advocate, Adam Smith, made it clear that government-imposed rules were essential if there was to be effective operation of free markets. It is all too often conveniently forgotten that Smith's "invisible hand" needed ground rules in order to successfully operate.

As our ancestors came to grips with the new technologies of the mid-19th century, they embraced the common law precedents. The duty to deal became a foundational concept of the industrial era. The Pacific Telegraph Act of 1860 required telegraph companies to carry all message traffic in the order received.[42] The same non-discriminatory access was applied to railroads by the first federal regulatory agency, the Interstate Commerce Commission (ICC). The dominant networks that defined the industrial revolution thus had the same kind of open access, duty to deal obligations as did medieval innkeepers and ferrymen. When the telephone network came along, it too had a duty to deal.

The 2015 Open Internet Rules of the Federal Communications Commission (FCC) extended the centuries old common law concept to the last mile internet service providers. While, unsurprisingly, the networks opposed such openness, the digital platform companies argued that such non-discriminatory access to the fundamental asset of a broadband connection to consumers was essential and should be open.

For 600 years the simple, yet irrefutable concept that the proprietor of a fundamental asset has a duty to make it available has stood the test of time and technology to remain valid today. While digital technology has redesigned the nature of bottlenecks, nothing has repealed the incentive behind the creation of such bottlenecks, nor the public interest remedy to their abuses.

For Networks the duty to deal means first come, first served nondiscriminatory access. The ferryman's duty to carry something across the river is no different than the telegraph, railroad, and telephone networks' duty to carry all comers indiscriminately. The concept promoted by today's internet service

providers that "digital is different" flies in the face of that history: network bottlenecks may enrich the perpetrator, but they certainly do not serve the public interest. Arguments that a duty to deal inhibits investment have been raised with every new technology and repeatedly proven self-servingly specious.

For Platforms the duty to deal means the inability to hoard a fundamental asset to the detriment of society. The medieval innkeeper was not required to feed travelers for free, but he was required not to withhold the sustenance he had collected and prepared. The innkeepers of the internet era are the platform companies that collect, aggregate and allocate digital information; like their analog predecessors, they are free to profit from their services, but the services must be openly available.

"Permissionless innovation" is a brilliant public relations moniker. It conjures up visions of faceless, small-minded bureaucrats deciding upon on the innovations of visionary entrepreneurs in a garage. It is, of course, a fictitious construction.

"Permissionless innovation" was made possible by rules

The companies benefitting from the digital revolution have sowed the idea that rules would hold back their magic. The secret behind their wonders is "permissionless innovation," they argue. It becomes the mantra whenever government oversight is proposed. The magic conjured by these innovators would somehow be broken if digital networks and platforms had to comply with regulatory oversight.

"Permissionless innovation" is a brilliant public relations moniker. It conjures up visions of faceless, small-minded bureaucrats deciding upon on the innovations of visionary entrepreneurs in a garage. It is, of course, a fictitious construction.

It was rules that allowed these companies to exist in the first place. The digital companies are sticklers for adherence to the technical standards of the internet. Without those rules, the low-cost collection of personal information would not be possible. Without those rules, the low-cost distribution of new digital creations would not exist.

While rules may be essential to protecting the backbone operations of the digital companies, rules suddenly became innovation killers when it comes to protecting the rights of the companies' consumers or the benefits of competition.

Applying the common law concepts of a "duty of care" and a "duty to deal" is to establish corporate responsibilities, not permission bureaucracies.

There is no governmental "permission" required to collect personal information. There should, however, be a responsibility to the consumer whose information is being collected and exploited. It is a "duty of care" about the effects of those activities on the wellbeing of the individual consumer and the public at large.

There is no governmental "permission" required for networks to offer new and expanded services to consumers. There should, however, be a responsibility for the modern ferry to the internet to serve everyone. It is a "duty to

deal" in a just and reasonable manner.

In the late 19th and early 20th centuries, as industrial capitalism took hold of the economy, it became important to protect capitalism by protecting consumers, workers and competitive markets. Now, in the early 21st century, it has become necessary to similarly protect internet capitalism, consumers, workers and competition.

For over six centuries simple, direct and decipherable common law concepts have governed economic affairs. The transformation of those economic affairs from analog activities to digital activities has not transformed the responsibilities of those who harness the new digital technology to provide new services.

Endnotes

1. https://loc.gov/about/fascinating-facts/
2. In item 3 below IDC reports 16.1 zettabytes of data generated in 2016 (or 16,100,00,00 terabytes) an amount that equates to 44,109,589 TB per day. The Library of Congress estimates it holds about 15 TB of data (https://blogs.loc.gov/thesignal/2012/04/a-library-of-congress-worth-of-data-its-all-in-how-you-define-it/) . 44,109,589 TB per day/15 TB = 2,940,639 Libraries of Congress
3. "Data Age 2025," International Data Corporation white paper, April 2017, https://www.seagate.com/www-content/our-story/trends/files/Seagate-WP-DataAge2025-March-2017.pdf - NOTE: the zettabyte measurement used in the chart's Y-axis is 10 followed by 21 zeroes
4. https://www.theverge.com/2015/6/2/8714345/tim-cook-epic-award-privacy-security
5. https://www.businessinsider.com/eric-schmidt-googles-policy-is-to-get-right-up-to-the-creepy-line-and-not-cross-it-2010-10
6. Axios, "Apple CEO calls for more privacy regulation, October 3, 2018
7. https://en.wikipedia.org/wiki/Relational_model
8. https://www.buzzfeednews.com/article/nicolenguyen/here-are-18-things-you-might-not-have-realized-facebook
9. For a further discussion, see Sarah E. Igo, *The Known Citizen: The History of Privacy in Modern America,* Harvard University Press, 2018, pp. 362-363.
10. Geoffrey A. Fowler, "No, Mark Zuckerberg, we're not really in control of our data," *The Washington Post,* April 12, 2018, https://washingtonpost.com/news/the-switch/wp/2018/04/12/no-mark-zuckerberg-were-not-really-in-control-of-our-data/
11. https://apnews.com/828aefab64d4411bac257a07c1af0ecb
12. https://www.theatlantic.com/technology/archive/2012/03/reading-the-privacy-policies-you-encounter-in-a-year-would-take-76-work-days/253851/
13. https://www.theatlantic.com/politics/archive/2018/10/mark-warner-is-coming-for-techs-too-powerful/572695/
14. https://www.nature.com/articles/srep01376
15. Edmond Locard, *Traité de Criminalistique,* J. Desvigne et ses fils, Lyon, 1931
16. Douglas C. Schmidt, *Google Data Collection,* Digital Content Next, https://digitalcontentnext.org/blog/2018/08/21/google-data-collection-research/
17. https://www.consumerreports.org/privacy/americans-want-more-say-in-privacy-of-personal-data/
18. https://www.mediapost.com/publications/article/324980/for-most-americans-personal-data-privacy-now-riva.html
19. https://www.breitbart.com/london/2018/03/23/bannon-governments-debase-citizenship-central-banks-debase-currency-state-capitalist-tech-firms-debase-personhood/
20. https://www.hollywoodreporter.com/news/at-t-ceo-wants-congress-create-privacy-rules-road-1148723
21. https://www.consumerreports.org/privacy/americans-want-more-say-in-privacy-of-personal-data/
22. https://www.nytimes.com/2018/04/11/business/zuckerberg-facebook-congress.html
23. https://www.blog.google/outreach-initiatives/public-policy/proposing-framework-data-protection-legislation/
24. https://www.ipc.on.ca/wp-content/uploads/Resources/7foundationalprinciples.pdf
25. *The Economist,* Special Report, "The ins and outs," June 30, 2018
26. https://www.capgemini.com/consulting/wp-content/uploads/sites/30/2017/07/digital_disruption_1.pdf
27. https://en.wikipedia.org/wiki/List_of_public_corporations_by_market_capitalization#2018
28. https://slate.com/technology/2018/02/apples-app-store-is-becoming-a-major-revenue-source-for-apple.html
29. https://www.statista.com/statistics/267161/market-share-of-search-engines-in-the-united-states/
30. https://www.statista.com/statistics/266136/global-market-share-held-by-smartphone-operating-systems/
31. https://www.statista.com/statistics/267161/market-share-of-search-engines-in-the-united-states/
32. https://www.nytimes.com/2018/09/07/technology/monopoly-antitrust-lina-khan-amazon.html
33. https://thenextweb.com/apps/2017/04/18/facebook-downloaded-app-netflix/
34. http://www.tvnewscheck.com/article/113380/google-to-dominate-local-digital-ads-in-2018
35. https://gigaom.com/2012/04/09/here-is-why-did-facebook-bought-instagram/
36. https://marketingland.com/report-facebook-takes-a-back-seat-to-instagram-as-ad-spend-on-the-facebook-owned-app-grows-177-244893

37. https://www.nytimes.com/2015/11/19/technology/how-amazons-long-game-yielded-a-retail-juggernaut.html
38. Tim Wu, *The Master Switch: The Rise and Fall of Information Empires,* Alfred A. Knopf, 2010, pp. 104-106
39. Jon Gertner, *The Idea Factory: Bel Labs and the Great Age of American Innovation,* The Penguin Press, 2012, p. 182
40. https://voxeu.org/article/how-antitrust-enforcement-can-spur-innovation
41. https://www.which.co.uk/money/banking/switching-your-bank/open-banking-sharing-your-financial-data-anscq4g8p62h
42. There was an exception for government messages.