Insight

Why A Data Portability Act Might Not Be An Effective Policy Path

WILL RINEHART | FEBRUARY 6, 2018

Voices in technology and policy are arguing that the largest tech companies, such as Google, Apple, Facebook, and Amazon, have too much power and face too little competition. Some have started calling for more direct regulation of these big companies as a way to curb their power. Others, who are more reluctant to regulate directly, have come up with a different idea for how to increase competition: data portability. This idea—simply put, making all platform data easily exportable—is gaining momentum as an alternative to increased regulation, as some believe it will provide the spark for platform competition as well as give ownership rights back to users.

Most of the proposals for creating data portability, however, miss what makes data valuable, and thus what gives these companies such power. Data isn't the key to Facebook, Amazon, and Google. Rather, it is the structure and processing tools around the data that drive these platforms. Further, the United States already tried an open-access regime to foster competition in the telecom sector, with lackluster results. Contrary to the standard bearers of this idea, there are very good reasons to think that regulations to create data portability won't be effective at all.

What Is Data Portability?

In 2016, the Obama Administration put out a request for information on the idea of data portability, asking for potential policy avenues. As the request noted, "if consumers cannot switch easily between platforms, it may be difficult for would-be services to enter the market, resulting in less innovation or higher prices." With data portability, the proposal argued, businesses would be induced to "compete with one another to offer better prices and higher quality services to win or retain a customer's business." This request in part reflected movement from the European Union: Next year, the EU's General Data Protection Regulation will go into effect and will include the right to data portability.

AMERICANACTIONFORUM.ORG

One specific proposal that has gained momentum is the "Social Graph Portability Act," an idea that comes from the University of Chicago's Luigi Zingales and Guy Rolnik. As a response to the "increasing monopolization of the technology industry," these economics professors want to reallocate "property rights via legislation to provide more incentives to compete." In their estimation, the idea is akin to telephone number portability, which "makes it easier to switch carriers, fostering competition by other carriers and reducing prices for consumers." As the authors explain,

> The same is possible in the social network space. It is sufficient to reassign to each customer the ownership of all the digital connections that she creates — what is known as a "social graph." If we owned our own social graph, we could sign into a Facebook competitor — call it MyBook — and, through that network, instantly reroute all our Facebook friends' messages to MyBook, as we reroute a phone call.

Their proposal, with its intuitive logic, puts Facebook, Google, Twitter and other platforms in the crosshairs. If users can reach their Facebook friends through a different social network, then they can try new social networks, and restore the benefit of competition. This kind of portability is not a new concept in the tech space. A similar kind of thinking was applied to AOL's Messenger (AIM) product as part of a condition of the merger with Time Warner. It was a precaution to guard "against AOL's ability to leverage its existing dominance in current IM into the broadband IM marketplace" in the words of FCC Chair Bill Kennard.

Data Isn't the Same as Processed Data

Underlying the "Social Graph Portability Act" and other versions of data portability is the belief that "the data" is the key to Facebook's ubiquity. For proponents of the data portability, user data is interchangeable with the social graph, a term that typically means the personal relationships of users mapped out. These proponents want you to be able to access this web of relationships from a range of platforms, much as we can call phone numbers from any carrier.

Contrary to the claims of portability proponents, however, it isn't data that gives Facebook power. The power rather lies in how this data is structured, processed, and contextualized. What Mark Zuckerburg has called the social graph is really his term for a much larger suite of interrelated databases and software tools that helps to analyze and understand the collected data. Some proclaim data to be the "new oil," but this analogy falls short, as a brief exploration of scalability demonstrates.

Scalability is simply the ability of programs to handle usage growth. The biggest players in

the space are united in their ability to innovate to solve problems *at scale* without interruptions, and it is hard to impress just how difficult of a problem engineers face. How Twitter, Facebook, eBay, YouTube, and Amazon have each solved various engineering issues at scaling each serves as a case study in its own right. Everywhere in tech, scaling is a concern and is taken seriously. The biggest tech players hold a conference dedicated to it, experts offering their services to upstarts help push past these barriers, and blogs that detail all of the cases are routinely cited. And yet, in tech regulation, the term doesn't even register.

Facebook's technology stack, the suite of technologies that it uses behind the scenes, clearly shows the importance of scaling, as much of the architecture was developed in-house to address the unique problems facing Facebook's vast troves of data. Facebook created BigPipe to dynamically serve pages faster, Haystack to efficiently store billions of photos, Unicorn for searching the social graph, TAO for storing graph information, Peregrine for querying, and MysteryMachine to help with end-to-end performance analysis. Nearly all of this design is open for others to use, and has been a significant boon to programmers in the ecosystem. The company also invested billions in content delivery networks to quickly deliver video, and it split the cost of an undersea cable with Microsoft to speed up information travel.

The vast investment that Facebook has put into programs for understanding and processing its users' data points to the fundamental flaw in the argument for data portability. Consider again Zingales and Rolnik's reasoning behind the Social Graph Portability Act. "Today Facebook provides developers with application-program interfaces that give them access to its customers' social graph, Facebook Connect and Graph A.P.I. Facebook controls these gates, retaining the right to cut off any developer who poses a competitive threat."

It is a fundamental misread to think that Facebook's control of its API—or application programming interface, the portal that allows different software components to communicate—is the source of Facebook's competitive advantage. In building for scale, Facebook and other social media sites have developed many interconnected computational assets, and it is the sum total of these pieces that makes them competitive, not simply one piece. As the official description from Facebook explains, the Graph API is "a low-level HTTP-based API," meaning the API does allow for a retrieval of information from the graph. And, to portability proponents' credit, Facebook does control its API. But when Mark Zuckerberg and others talk about the social graph, they are describing the insights derived from the social graph in conjunction with the rest of the resources of the network.

The demise of MySpace demonstrates the competitive advantage that Facebook derives from its programs that analyze data at scale. Nearly a decade later it is difficult to

remember, but MySpace lost out to Facebook because it hit a wall with its technology and slogged through upgrades. Web site error rates were as high as 30 or 40 percent at the time because MySpace wasn't created with the kind of systematic approach to computer engineering that went into Yahoo, eBay, Google, and Facebook. In contrast, those sites had at most 1 percent error rates. Facebook's ability to capitalize on the data it had allowed it to triumph competitively against MySpace.

Data isn't the most important asset in the digital economy, but the talent needed to make that data useful is crucial. Herein lies the flaw with the "new oil" analogy. Saudi Arabia has a tremendously valuable asset in the volume of its oil because the ability to process oil is well known and relatively simple. Not so with data, however: Data is ubiquitous, but what isn't widespread is understanding what kind of information you might have and doing something useful with it. Requiring data portability does little to deal with the very real challenges that face the competitors of Facebook, Amazon, and Google. Entrants cannot merely compete by collecting the same kind of data. They need to build better sets of tools to understand information and make it useful for consumers.

Telecom Unbundling and Lower Investment

Reassigning rights to data misses the fundamental point about data. It is not the mere presence of data that confers an advantage. Tech companies have to invest in their platforms, like any other asset. Instead of likening data to telephone numbers, the better analogy to understand how data portability would work comes from a different part of the telecom sector.

In the 1996 Telecommunications Act, Congress attempted to increase competition in local telecommunications markets by unbundling network elements of the telephone system. Unbundling is a regulatory regime that allows competitors of a telecommunications operator to use those incumbents' network assets, such as telephone lines or switches, at a regulated rate. In creating a space for competitors, the designers of the regime hoped the local telecommunications markets would become more competitive, with competitors eventually building their own networks to compete directly with the incumbents.

In practice, the outcomes have been far less than stellar. For one, pricing unbundled elements is an inherently difficult task, and pricing tensions plagued telecom unbundling nearly everywhere it has been tried. In the United States, United Kingdom, New Zealand, Canada, and Germany, all the primary goals of managed competition were found to be wanting. Retail prices in the United States were slightly higher after the change, but investment was significantly down for the incumbents. Thus, this regime didn't yield lower prices as some had hoped.

Moreover, unbundling wasn't found to be connected to future investment. The hope was that competitors would build their own networks to compete with the biggest players, also known as the stepping-stone hypothesis. What happened was very different, however. Competitors were allowed to lease network elements during periods of profitability but could return them during downturns. Thus, when demand increased, competitors could also ramp up their networks and when prices fell, they had the ability to dump those assets. The incumbents thus bore the burden on both ends, and faced incentives not to invest, which seems to have happened in many countries.

None of the unbundling experience has been applied to the debate regarding data portability, yet it raises big questions. Would the proposed Act give competitors access to every element of Facebook, Amazon, or Google's platform without any prices? What about content delivery networks and data centers, which costs billions of dollars to develop? Will those also be unpriced? Indeed, throughout the discussion, it is continually said that data is owned by the user with little discussion of the costs incurred by the platforms to collect and understand data. Since data portability advocates understand this move as merely reassigning rights, they are blind to the "enormous investment of time by some of the world's most highly compensated engineers" needed to bring these tools into use.

Indeed, if regulations were used to effectively open up Facebook and other platforms, then the network elements on which these networks rely would face many of the same problems. The incumbents would bear a burden on both ends and be pressed not to invest.

New Environment, Old Ideas

The fundamental ideas behind data portability have been tried and found wanting. When AOL was forced to make its users' data portable, the expected competitive leverage never happened and the services that were supposed to sprout from AIM never materialized. Fundamentally, the FCC misassessed where the power lay in AIM, assuming it came from AOL's wide user base. Instead, AIM provided a unique service that took nearly another decade to replicate. Then as now, the existence of a social network and the accompanying data doesn't confer an unassailable advantage. What does confer an advantage is a suite of useful services.

AIM nevertheless failed, and the reasons here are also instructive. AOL often kept the service at arm's length and never dedicated the kind of resources needed to keep the service competitive. AOL also shunned outside developers from using AIM as a platform, which limited its use. In short, AIM was overtaken due to neglect. Because it didn't innovate, develop its network, and continue to roll out new features, AIM was retired last year. The same dynamic applies to the largest tech companies today—even with the vast

---

amounts of data they have, unless they invest and innovate, they will struggle.

Advocates and wonks pressing for regulation like graph portability need to dive deeper, understand where value is located, and deal with the specifics. The competitive edge of large social sites and platforms isn't based just in the data, but in how that data is organized and processed. The lessons of the telecom unbundling regime are also worth heeding: An effective data portability regime would have to open up all of a company's tools for processing data, which could reduce their incentives to invest. As it stands, the Social Graph Portability Act proposal doesn't take this subject seriously, and that is worrying.