2021

Taking It With You: Platform Barriers to Entry and the Limits of Data Portability

Gabriel Nicholas
New York University School of Law

TAKING IT WITH YOU: PLATFORM BARRIERS TO ENTRY AND THE LIMITS OF DATA PORTABILITY

Gabriel Nicholas*

ABSTRACT

Policymakers are faced with a vexing problem: how to increase competition in a tech sector dominated by a few giants. One answer proposed and adopted by regulators in the United States and abroad is to require large platforms to allow consumers to move their data from one platform to another, an approach known as data portability. Facebook, Google, Apple, and other major tech companies have enthusiastically supported data portability through their own technical and political initiatives. Today, data portability has taken hold as one of the go-to solutions to address the tech industry's competition concerns.

This Article argues that despite the regulatory and industry alliance around data portability, today's public and private data portability efforts are unlikely to meaningfully improve competition. This is because current portability efforts focus solely on mitigating switching costs, ignoring other barriers to entry that may preclude new platforms from entering the market. The technical implementations of data portability encouraged by existing regulation—namely one-off exports and API interoperability—address switching costs but not the barriers of network effects, unique data access, and economies of scale. This Article proposes a new approach to better alleviate these other barriers called collective portability, which would allow groups of users to coordinate to transfer data they share to a new platform, all at once. Although not a panacea, collective portability would provide a meaningful alternative to existing approaches while avoiding both the privacy/competitive utility trade off of one-off exports and the hard-to-regulate power dynamics of APIs.

INTRODUCTION

On March 30, 2019, in response to fomenting anti-Facebook sentiment, Mark Zuckerberg released an op-ed titled, "Four Ideas to Regulate the Internet."[1] His first three ideas regarded speech, political advertisements, and privacy. The fourth addressed concerns that Facebook and its ilk had grown too large for any other platform to compete.[2] Zuckerberg called for regula-

---

1.   Mark Zuckerberg, *The Internet Needs New Rules, Let's Start in These Four Areas*, WASH. POST (Mar. 30, 2019, 3:00 PM), https://www.washingtonpost.com/opinions/mark-zuckerberg-the-internet-needs-new-rules-lets-start-in-these-four-areas/2019/03/29/9e6f0504-521a-11e9-a3f7-78b7525a8d5f_story.html.

2.   *See, e.g.*, Lina Khan, *Amazon's Antitrust Paradox*, 126 YALE L.J. 710 (2017); Robert Reich, *Break up Facebook (and While We're at It, Google, Apple and Amazon)*, GUARDIAN (Nov. 20, 2018, 3:00 AM), https://www.theguardian.com/commentisfree/2018/nov/20/facebook-google-antitrust-laws-gilded-age; Joe Nocera, *Easiest Fix for Facebook: Break It up*, BLOOMBERG (Nov. 21, 2018, 11:44 AM), https://www.bloomberg.com/opinion/articles/2018-11-21/fix-facebook-by-breaking-it-up-under-antitrust-regulation.

tion to guarantee the principal of data portability, noting that, "[i]f you share data with one service, you should be able to move it to another. This gives people choice and enables developers to innovate and compete."[3]

Zuckerberg's op-ed signaled a strategic shift in Facebook's public policy strategy. Instead of fighting to keep Facebook data within the walls of the platform,[4] Facebook would allow for "data portability"—the ability for users to move their data from one service to another. It was an instrument of self-regulation that responded directly to mounting calls for antitrust action.[5] In the same way that the free movement of capital allows for competitive markets, so too would the free movement of data allow for competitive platforms, at least in theory.[6] Facebook was not alone in adopting this approach—Apple, Google, Microsoft, and Twitter all included data portability in their competition policy efforts.[7] Data portability was also part of the ACCESS Act,[8] introduced by Senators Warner, Hawley, and Blumenthal seven months after Zuckerberg's op-ed, to improve competition in the tech sector.[9]

Although data portability laws already exist in California, the European Union, and Singapore,[10] they are not primarily aimed at improving tech sec-

---

3.   Zuckerberg, *supra* note 1.

4.   *See generally* Mark Andrejevic, *Privacy, Exploitation, and the Digital Enclosure*, 1 AMSTERDAM L.F. 47 (2009); Kevin Bankston, *How We Can 'Free' Our Facebook Friends*, NEW AMERICA (June 28, 2018), https://www.newamerica.org/weekly/edition-211/how-we-can-free-our-facebook-friends.

5.   Recent antitrust stirrings indicate that the government is interested in increasing competition among tech companies, particularly Alphabet, Amazon, Facebook, and Apple, which are all facing antitrust scrutiny by the Justice Department and other federal and state organizations. *See* Daisuke Wakabayashi et al., *16 Ways Facebook, Google, Apple and Amazon Are in Government Cross Hairs*, N.Y. TIMES (Sept. 9, 2019), https://www.nytimes.com/interactive/2019/technology/tech-investigations.html.

6.   *See generally* VICTOR MAYER-SCHÖNBERGER & THOMAS RAMGE, REINVENTING CAPITALISM IN THE AGE OF BIG DATA (2018); Peter Swire, *Markets, Self-Regulation, and Government Enforcement in the Protection of Personal Information, in Privacy and Self-Regulation in the Information Age*, DEP'T OF COM. (June 1997), https://papers.ssrn.com/abstract=11472.

7.   The Data Transfer Project is a joint framework intended to make it easier for companies to let users port their data between. It is an open source project being led by Apple, Facebook, Google, Microsoft, and Twitter. *See generally Data Transfer Project Overview and Fundamentals*, DATA TRANSFER PROJECT (July 20, 2018), https://datatransferproject.dev/dtp-overview.pdf.

8.   Augmenting Compatibility and Competition by Enabling Service Switching Act of 2019, S. 2658, 116th Cong. (2019).

9.   *See also* Consumer Online Privacy Rights Act, S. 2986, 116th Cong. § 105(a) (2019) (providing for a right to data portability stronger than that in the CCPA because it requires data to be made available in a computer-readable format, but weaker than the right provided in the GDPR because it does not require data controllers to transfer the data to a new entity upon request).

10.   CAL. CIV. CODE §§ 1798.100-1799.100; Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Re-

tor competition. The California Consumer Privacy Act (CCPA), for example, adopts portability to allow consumers to hold platforms accountable for how they track and sell personal data.[11] The law offers little utility to new market entrants since it does not even require platforms to make their data available in a machine-readable format—the CCPA explicitly allows companies to transfer data via a toll-free number.[12]

The right to data portability in Europe's General Data Protection Regulation (GDPR)[13] offers more utility to competitors, but not as much as it could, given that its main purpose is to grant consumers greater control over their own data.[14] Article 20 requires data controllers to allow consumers to export their personal data in a machine-readable format, and where technically feasible, directly transmit data to the new controller.[15] European legal scholars argue that this improves competition by lowering the cost for users to change platforms.[16] But in certain circumstances, the aims of data control and competition directly conflict.[17] For example, if a data subject tries to port data that overlaps with someone else's, the GDPR states that it is more important to respect the privacy and erasure rights of the second data subject than portability rights of the first.[18]

The Lei Geral de Proteção de Dados (LGPD)—Brazil's data protection law—also gives data subjects the right to port their data from one data controller to another.[19] However, unlike the GDPR, which requires the consid-

---

pealing Directive 95/46/EC (General Data Protection Regulation), art. 20, 2016 O.J. (L 119) [hereinafter GDPR]; Personal Data Protection (Amendment) Act 2020 (Sing.).

11.   The CCPA gives consumers rights to access all personal data a platform has collected on them, delete that information, and access certain knowledge about how their personal data gets sold, including what data has been sold and what category of business has bought it. Users also have the right to prevent buyers from reselling it without permission. CAL. CIV. CODE §§ 1798.105, .110, .115.

12.   *See id.* § 1798.130. The CCPA also only requires platforms to send twelve months' worth of user data.

13.   GDPR. art. 20.

14.   *Id.* recital 68 ("To further strengthen the control over his or her own data, where the processing of personal data is carried out by automated means, the data subject should also be allowed to receive personal data concerning him or her which he or she has provided to a controller").

15.   *Id.* art. 20.

16.   *See generally* Inge Graef, *Blurring Boundaries of Consumer Welfare: How to Create Synergies Between Competition, Consumer and Data Protection Law in Digital Markets, in* PERSONAL DATA IN COMPETITION, CONSUMER PROTECTION AND INTELLECTUAL PROPERTY LAW (Mor Bakhoum et al., eds. 2018); Damien Geradin & Monika Kuschewsky, *Competition Law and Personal Data: Preliminary Thoughts on a Complex Issue* (Feb. 13, 2013) (unpublished manuscript) (https://ssrn.com/abstract=2216088).

17.   *See generally* Tal Zarsky, *The Privacy-Innovation Conundrum*, 19 LEWIS & CLARK L. REV. 115, 141 (2015).

18.   *See* GDPR, *supra* note 13, art. 20.

19.   Abigayle Erickson, Note, *Comparative Analysis of the EU's GDPR and Brazil's LGPD: Enforcement Challenges with the LGPD*, 44 BROOK. J. INT'L L. 862, 883 (2019); Lei No. 13.709, de 14 de Agosto de 2018, Diário Oficial Da União [D.O.U.] de 15.8.2018 (Braz.),

eration of the data subject's individual rights before the "legitimate interests" of the data controller, the LGPD allows controllers to weigh the two considerations against one another.[20] It is still unclear what this balance means for the competitive utility of Brazilian data portability, since the law has only recently gone into effect as of this writing.

Only some of these data portability laws attempt to address platform barriers to entry, and none attempt to address barriers beyond switching costs.[21] These barriers to entry, here meaning the structural economic disadvantages faced by some firms but not others,[22] may still preclude new platforms from entering the market. For example, a platform that can overcome an incumbent's switching costs may still fail to overcome its network effects or economies of scale.

In this Article, I argue that the technical approaches to data portability, as encouraged by existing regulation, fail to improve barriers to entry sufficiently to allow new platforms to enter the market. As an alternative to the current popular approaches to implementing portability, namely one-off exports and APIs, I recommend a new approach that I call *collective portability* to better address these barriers. Collective portability allows groups of users to coordinate in order to transfer data they share to a new platform. This not only helps new market entrants overcome switching costs, but also addresses the barriers of unique data access, economies of scale, and network effects.

In making suggestions on how to evaluate and improve data portability's viability as a tool for regulating competition, this Article seeks to contribute to two broader bodies of literature. The first is the literature on data as an economic asset. Scholars often discuss the economic advantages and normative drawbacks of data-rich technical systems, but rarely discuss the contexts in which data-based advantages fail to be realized. To the extent this conversation occurs, it is mostly done through analogies, such as de-

---

translated in BRAZILIAN GENERAL DATA PROTECTION LAW (Ronaldo Lemos et al. trans., 2019), https://iapp.org/media/pdf/resource_center/Brazilian_General_Data_Protection_ Law.pdf (last visited June 7, 2020) [hereinafter LGPD].

20.   *See* Steven Blickensderfer et al., *Brazil's New Data Protection Law: An Overview and Four Key Takeaways for U.S. Companies*, NAT'L L. REV. (May 2, 2019), https:/ /www.natlawreview.com/article/brazil-s-new-data-protection-law-overview-and-four-key- takeaways-us-companies; Richie Koch, *What is the LGPD? Brazil's Version of the GDPR*, https://gdpr.eu/gdpr-vs-lgpd (last visited Dec. 22, 2020).

21.   *See* Peter Swire & Yianni Lagos, *Why the Right to Data Portability Likely Reduces Consumer Welfare*, 72 MD. L. REV. 338, 349 (2013).

22.   *See, e.g.*, ORG. FOR ECON. CO-OPERATION AND DEV., GLOSSARY OF INDUSTRIAL ORGANISATION ECONOMICS AND COMPETITION LAW 13–14 (1990), http://www.oecd.org /regreform/sectors/2376087.pdf; Yafit Lev-Aretz & Katherine J. Strandburg, *Regulation and Innovation: Approaching Market Failure from Both Sides* 4 (Pub. L. & Legal Theory Rsch. Paper Series, Working Paper No. 19-48, 2019), https://ssrn.com/abstract=3462522.

bates over whether data is the new oil.[23] Through my analysis of data portability, I show how data is not a fungible asset with value in isolation. Rather, the value of data depends on its context, and the transferability of that value depends on the capacity to maintain that context between services.

Second, this Article seeks to contribute to the larger literature on regulating tech companies. As scholars and policymakers consider how to manage platform giants, data portability will be an important tool in the regulatory tool belt. But data portability is not a panacea, and what works as a matter of policy may fail in technical implementation. This Article examines various real-world examples of data portability that are well-known in Silicon Valley but not often analyzed in legal scholarship. They will hopefully serve as a good prism through which to interpret existing scholarship and bolster new scholarship.

I build my argument on three types of sources. The first is theoretical work done on barriers to entry from the fields of law, economics, and information systems. Second is anecdotes from real data portability regimes and analogous efforts, mostly from real tech companies. Despite the insistence that data portability improves competitiveness in the tech sector, I could not find any products built off of data specifically from one-off exports[24] (despite the fact that Facebook has allowed users to download their data since 2010[25] and Google since 2011[26]). Therefore, for this part of the analysis, I turn to analogous initiatives in the telecom and banking sectors. The third type of evidence is real ported data. I downloaded and analyzed my own personal data from seven different platforms—Facebook, Google,

---

23.   *See* Lauren Henry Scholz, *Big Data Is Not Big Oil: The Role of Analogy in the Law of New Technologies*, 86 TENN. L. REV. 863 (2020). *But see* Jonathan Vanian, *Why Data Is the New Oil*, FORTUNE (July 11, 2016), https://fortune.com/2016/07/11/data-oil-brainstorm-tech; Joris Toonders, *Data is the New Oil of the Digital Economy*, WIRED, http://www.wired.com/2014/07/data-new-oil-digital-economy (last visited Feb. 11, 2021).

24.   It is technically possible for a market entrant to build a new product off of a one-off export since they are computer readable, but in practice they are difficult if not impossible for developers to integrate into new or existing products. Data senders tend to provide little to no documentation about the structure of data users get from an export, leaving receivers to reverse engineer it themselves. Of the seven data portability exports studied for this Article, six had no documentation. The seventh, Twitter, seems to use something similar to its own documented, external facing API, though this is not confirmed anywhere. Also, the Twitter data set was the only one not in a cleanly formatted JSON. None of them had any versioning, which meant that these platforms could introduce breaking changes at any time. *See* GABRIEL NICHOLAS & MICHAEL WEINBERG, ENGELBERG CTR. ON INNOVATION L. & POL'Y, DATA PORTABILITY AND PLATFORM COMPETITION: IS USER DATA EXPORTED FROM FACEBOOK ACTUALLY USEFUL TO COMPETITORS? 21 (Nov. 2019).

25.   Alexia Tsotsis, *Facebook Now Allows You to "Download Your Information"*, TECHCRUNCH (Oct. 6, 2010, 1:52 PM), https://techcrunch.com/2010/10/06/facebook-now-allows-you-to-download-your-information.

26.   *See* GOOGLE, *The Data Liberation Front Delivers Google Takeout*, DATA LIBERATION BLOG (June 28, 2011, 1:51 PM), http://dataliberation.blogspot.com/2011/06/data-liberation-front-delivers-google.html.

Apple's iCloud, Instagram, Snapchat, Spotify, and Twitter to understand how platforms can and cannot leverage ported data to compete.

Part I defines "data portability" and delineates between two popular approaches, one-off exports and APIs. Part II assesses how effectively each of these approaches helps competitors overcome four different barriers to entry raised by big data: switching costs, network effects, economies of scale, and unique data access. Part III highlights additional limitations of API portability, based on the market behavior of users and platforms, and opportunities, based on its innovative potential. I then propose a new approach in Part IV—*collective portability*. I review it through the lens of Facebook's Download Your Information (DYI) tool, and then explore questions, hypotheticals, and challenges to the approach. I conclude with recommendations for regulators.

I. Current Approaches to Data Portability

The International Organization for Standardization defines "data portability" as the "ability to easily transfer data from one system to another without being required to re-enter data".[27] This definition broadly encompasses any means of "easy" transfer[28] for any type of data. But data portability laws tend to have a narrower focus when defining data portability. The GDPR focuses its definition of data portability on personal data—Article 20 gives Europeans the rights to obtain "personal data concerning [a data subject] which he or she has provided."[29] It does not include information a data controller inferred from provided personal data or any non-personal data, since that data is not considered "provided" by the data subject.[30]

Data portability laws also do not provide a data subject access to someone else's personal data, even if it overlaps with her own. This is particularly relevant in the context of deletion laws, like the GDPR's "right to be forgotten."[31] If Alice is tagged in a photo that Bob took and uploaded, for example, portability regimes under the GDPR may allow Bob to port that photo but not Alice, since the photo is Bob's data and someday he may want to delete it, which he could not do if Alice had downloaded it.[32]

---

27.   Int'l Standards Org. [ISO], *Automation Systems and Integration—Oil and Gas Interoperability*, at 3.23, ISO Standard No. 18101-1:2019 (2019), https://www.iso.org/obp/ui/#iso:std:iso:ts:18101:-1:ed-1:v1:en:term:3.23 [hereinafter ISO].

28.   Forms of transfer that would not be considered easy include manual data entry, an external tool, or having a specialist do it. *See infra* Section II.A.

29.   GDPR, *supra* note 13, art. 20.

30.   This Article will largely focus on personal-information-based products and services since these are the ones already subject to regulation.

31.   GDPR, *supra* note 13, recital 65.

32.   When exactly the "right to be forgotten" applies to social media data has somewhat yet to be determined. Many cases fall between the gaps of the extreme examples given by the European Commission. *See generally European Data Commission, Do We Always Have to*

By and large though, platforms can themselves decide which user data they make portable. To better conceptualize this, imagine a platform represented as a circle-and-line graph of entities and connections, or *points* and *edges*.[33] Each point represents a person or any discrete piece of content, like a message or a video. Each edge represents a relationship between two points, like between two people (e.g., friends on Facebook), a person and content (e.g., liking a photo) or two pieces of content (e.g., a comment on a photo). A specific data portability regime can be visualized as which edges and points an individual user is allowed to transfer, and in what context. A software service may not make all points and edges portable, particularly when they do not fall clearly under the control of a single user, and data may fall through the gaps.

There are two main types of portability, and they differ in how they allow data to move. The first is *one-off export portability*, whereby users can download a snapshot of the data they have on one platform in a form that can be uploaded to another.[34] One-off exports do not require the platform which data is being transferred away from (the *data sender*) and the platform which data is being transferred into (the *data receiver*) to have a direct connection.[35] Often, the user acts as an intermediary between the two platforms, moving data from one to the other. This gives the user full control over which data he sends, and it allows him to delete his data from the sending platform if he so desires. In practice, one-off exports often take a long time for data controllers to prepare—minutes to hours on some platforms,[36]

Delete Personal Data if a Person Asks?, Eur. Comm'n, https://ec.europa.eu/info/law-topic/data-protection/reform/rules-business-and-organisations/dealing-citizens/do-we-always-have-delete-personal-data-if-person-asks_en (last visited Feb. 13, 2021). The European Data Protection Board has begun to fill in some of these gaps by differentiating between "delisting" (from search engines) and "full erasure". *See Guidelines 5/2019 on the Criteria of the Right to Be Forgotten in the Search Engines Cases Under the GDPR*, Eur. Data Protection Bd. (Dec. 2, 2019), https://edpb.europa.eu/sites/edpb/files/consultation/edpb_guidelines_201905_rtbfsearchengines_forpublicconsultation.pdf.

33.	*See* Alan Gibbons, Algorithmic Graph Theory 1 (1985).

34.	The terminology can be confusing, because some refer to one-off export portability as "data portability" and API portability as "interoperability". *See generally* Oscar Borgogno & Giuseppe Colangelo, *Data Sharing and Interoperability: Fostering Innovation and Competition Through APIs*, 35 Comput. L. & Sec. Rev. 1 (2019); *see generally* Paul De Hert et al., *The Right to Data Portability in the GDPR: Towards User-Centric Interoperability of Digital Services*, 34 Comput. L. & Sec. Rev. 193 (2018). I use the terms "one-off export portability" and "API portability" to avoid confusion and remind the reader that both, by the ISO definition, are forms of data portability.

35.	*See* Erin Egan, Facebook, Data Portability and Privacy 11 (2019), https://about.fb.com/wp-content/uploads/2020/02/data-portability-privacy-white-paper.pdf [hereinafter Facebook Portability White Paper]. Egan refers to this relationship as an "open transfer."

36.	*E.g.*, Facebook, Twitter.

days to weeks on others[37]—because the data sender may have to move a lot of data at once that is not stored in an easily accessible format.[38]

The second type of data portability is *interoperability*, whereby users allow two or more platforms to exchange information directly with one another. The ongoing connection between systems facilitates the transfer of live, up-to-date information, often through defined and documented protocols called application programming interfaces ("APIs"). To avoid confusion with uses of the term "interoperability" in other contexts,[39] I will refer to this approach as *API portability*.[40]

In contrast with one-off exports, APIs require the data sender and data receiver to maintain an ongoing relationship.[41] The receiver requests data from the sender, allowing it to obtain data faster and more frequently. It also lets the data sender observe and control what, when, and how the receiver gets data, usually through rate limiting (which limits how frequently it can make API requests) and key registration (which requires the receiver to register with the sender to access the API, thus allowing the sender to see how it uses the API and cut off access if necessary). There are reasons outside of competition for data senders to retain this kind of control, including security, preventing spam, and keeping costs down. However, these mechanisms have also been used for anti-competitive purposes, as will be discussed later.[42]

To understand the difference between one-off exports and APIs, imagine that Carlo keeps track of his recipes in an app called CookUp but wants to move them to a new app called Tazte. If CookUp allows for one-off export portability, Carlo can download his recipes, upload them to Tazte, and, if he wants, delete that data from CookUp. The export is just a snapshot of Carlo's data on CookUp, so if Carlo later adds a recipe to CookUp that same recipe would not also appear on Tazte. However, if CookUp gives API access to Tazte, perhaps through a "Sign In With CookUp" button, the two

---

37.   *E.g.*, Spotify, Snapchat.

38.   The databases used to store this data are often optimized for some operations over others. Since all platforms explored here were originally developed after portability laws existed, it is unlikely that their databases are optimized for downloading all of a user's data.

39.   "Interoperability" is often seen as synonymous with "compatibility". In this context, the ISO definition makes the most sense: the "ability of two or more systems or applications to exchange information and to mutually use the information that has been exchanged." Information technology—Cloud computing—Interoperability and portability, No. 19941 Int'l Standards Org. & Int'l Electrotechnical Comm'n. [ISO/IEC] (2017), https://www.iso.org/obp/ui/#iso:std:iso-iec:19941:ed-1:v1:en.

40.   APIs can be used for more than moving data around. They can also allow applications to expose features of their system (e.g., database storage, algorithms.) In this Article, I only consider the strict portability use case.

41.   FACEBOOK PORTABILITY WHITE PAPER, *supra* note 35, at 12. Facebook calls this a "partnership transfer."

42.   *See infra* Section III.A.

systems can directly exchange live data. Thus, if Carlo adds or deletes a recipe to CookUp, the two lists could potentially sync.

One-off export and API portability are not strictly dichotomous. For example, a data sender and data receiver may be in direct communication for a brief period of time and then be cut off from one another. Or a data sender may only let users transfer their data to a platform if certain conditions are met, such as only allowing health data to be sent to another HIPAA-compliant platform.[43] This will become relevant in Part IV when I propose a new approach to portability that mixes elements from both one-off and API portability.

II. Data Portability and Barriers to Entry

Competition problems in the tech sector differ meaningfully from competition problems in other sectors because technology firms often have business models that depend on and leverage the advantages of big data.[44] Scholars argue that disparities in data held by technology firms can create barriers to entry that make it difficult for non-dominant platforms to compete.[45] In this section, I review four of these barriers: switching costs, unique data access, economies of scale, and network effects. I have ordered them based on their ability to be alleviated by current approaches to data portability from best to worst. These four barriers will serve as an operational definition of competition[46] and will be used to evaluate the competitive utility of one-off and API portability respectively.

A. *Switching Costs*

If a user has a lot of her data on one platform, there may be a significant cost for her to switch to another platform. For example, if she has rated hundreds of books read on Amazon's Goodreads platform, she may not be willing to switch to a new, more preferable platform if it means having to

---

43.   Facebook Portability White Paper, *supra* note 35, at 12. Facebook calls this a "conditioned transfer."

44.   I define this term with the four Vs of big data: volume, velocity, variety, veracity. In other words, "big data" sets are large, frequently updated, gathered from various sources, and reasonably accurate.

45.   *See, e.g.*, *The Four V's of Big Data*, IBM Big Data & Analytics Hub (2015), www.ibmbigdatahub.com/infographic/four-vs-big-data. Different sources expand or contract this mnemonic, ranging between three and six V's. Matt Griffin et al., *A Case Study: Analyzing City Vitality with Four Pillars of Activity—Live, Work, Shop, and Play*, 4 Big Data 60, 61 (2016), https://www.liebertpub.com/doi/full/10.1089/big.2015.0043.

46.   Competition can be an elusive term so it is easier to evaluate by taking on a non-comprehensive but also non-recursive operationalization. *See generally* Paul McNulty, *Economic Theory and the Meaning of Competition*, 82 Q.J. Econ. 639, 639–40 (1968).

manually re-enter every book she read. Switching costs thereby act as a barrier to entry.[47]

One-off data exports are naturally suited to lower switching costs since data portability is by definition the "ability to easily transfer data."[48] A one-off export lowers switching costs in the Goodreads example because the user values her data (reading history) for its own sake, not just in the context of other data on the Goodreads platform that could not be exported (e.g., comments from other users). One-off exports work similarly well for other types of data that users value for its own sake but are difficult to move, such as photos, videos, and blog posts.[49]

There is historical precedent for one-off exports[50] successfully improving competition in the telecom sector. On November 24, 2003, the Federal Communications Commission (FCC) mandated mobile phone number portability, requiring all wireless carriers to allow customers who wished to switch to move their phone number to another carrier.[51] Previously, if a customer wanted to change carriers, he would also have to abandon his existing phone number. By removing this switching cost, the FCC hoped to increase competition between carriers.[52] Looking at cost as a proxy for competition, mobile number portability was an across the board success. Small, medium, and large phone plans all became cheaper, and price dispersion between companies offering analogous plans decreased.[53]

Well-designed APIs can drive switching costs even lower than one-off exports. Streamlined design elements like the "Sign in with Facebook" button make for a more seamless transfer experience than downloading and uploading an export file. APIs also allow users to change services without

---

47.   Daniel Rubinfeld & Michal Gal, *Access Barriers to Big Data,* 59 ARIZ. L. REV. 339, 364 (2017).

48.   ISO, *supra* note 27.

49.   Facebook has already taken this approach with its port to Google Photos project. *See* Steve Satterfield, *Driving Innovation in Data Portability with a New Photo Transfer Tool*, FACEBOOK (Dec. 2, 2019), https://about.fb.com/news/2019/12/data-portability-photo-transfer-tool.

50.   *See generally* Minjung Park, *The Economic Impact of Wireless Number Portability*, 59 J. INDUS. ECON. 714 (2011). Mobile phone number portability is an example of a one-off export, not interoperability. When transferring a number, the sender and receiver do need to communicate with one another once the relevant data, originally allocated by the North American Numbering Plan Administrator, is sent. The FCC also mandates this process happens in a timely manner, and it usually takes about two hours.

51.   *See generally* Telephone Number Portability/CTIA Petitions for Declaratory Ruling on Wireline-Wireless Porting Issues, *Memorandum Opinion and Order*, FCC 03-284, 18 FCC Rcd. 23697 (2003). This was actually only in the top 100 Metropolitan Statistical Areas. It only applied to all markets (despite a Petition for Forbearance filed by a number of telecoms) starting May 24, 2004.

52.   *Id.* at 4.

53.   For example, unlimited anytime minute plans became more similarly priced between AT&T and Verizon. Park, s*upra* note 50, at 734–36.

abandoning the original, further reducing both switching costs and the cost of returning to the incumbent.[54]

However, data portability does not necessarily cause data to flow from dominant services to competitors. The phone number portability example sidesteps this dynamic because the telecom industry mostly consists of a few dominant players. The tech industry, on the other hand, has a few large players and many small ones, and improving competition means making it easier for the two groups to compete on the same footing. Yet reducing switching costs might actually *harm* smaller competitors by helping incumbent platforms with strong network effects pull in users.

This dynamic plays out on customer relationship management platforms (CRMs). A small business may opt for a smaller, cheaper CRM, such as Pipedrive, that competes with the industry-standard CRM, Salesforce, on price.[55] Once the business grows though, they will have more clients who use Salesforce and will want to better integrate with their systems. The startup is incentivized to switch to the strongly networked Salesforce as soon as they are able. Currently, this can take a while because startups need to either buy a third-party tool or pay a specialist to switch CRMs.[56] Lower switching costs through data portability could cause businesses to switch away from small CRMs sooner than they otherwise would, making it even more difficult for them to compete with Salesforce's network effects.

B. *Unique Data Access*

Critics who argue that big data does not act as a barrier to entry point out that data is a non-rivalrous, non-exclusive resource.[57] However, there are cases in which a platform can have unique and functionally exclusive access to certain data. For example, an IoT thermometer company may have unique access to data concerning a patient's body temperature, or an oil company may have unique access to particular geothermal data.

Data portability can somewhat mitigate the advantage of unique data access. If an IoT thermometer company makes its data portable, a health tracking app can use that data to give its users a better sense of their overall health. However, the data sender will likely have more context about how the data was collected since it has greater knowledge of its own internal systems. This translates into knowledge about which data is more trustworthy.

---

54.   *See infra* Section III.B.

55.   Interview with Berk Birand, CEO, Fero Labs, in N.Y.C., N.Y. (Nov. 21, 2019); D.P. Taylor, *Pipedrive vs. Salesforce: A CRM Comparison for 2021*, BLUEPRINT (Jan. 6, 2021), https://www.fool.com/the-blueprint/pipedrive-vs-salesforce.

56.   *Id.*; *see generally* IMPORT2 WIZARD, https://www.import2.com.

57.   *See generally* Anja Lambrecht & Catherine Tucker, Can Big Data Protect a Firm from Competition? (Dec. 18, 2015) (unpublished manuscript) (https://ssrn.com /abstract=2705530); Darren Tucker & Hill Wellford, *Big Mistakes Regarding Big Data*, ANTITRUST SOURCE, A.B.A. (Dec. 1, 2014), https://ssrn.com/abstract=2549044.

Additionally, the data sender may withhold some of the data it collects. It may not want to overload the receiver with excessive data, or it may want withhold data for its own advantage. The IoT thermometer company may choose not to expose its internal diagnostic data because it is unnecessary, or because it undermines the confidence of the temperature taken, or because it undermines user privacy.

The U.K. Department for Business Innovation & Skills is attempting to lower the barriers of unique data access in the financial tech sector through a portability initiative called Midata. Midata makes it easier for consumers to access information about their bank transactions in hopes of allowing them to compare the prices of different products using their own data.[58] The initiative first focused on the banking industry though there are plans to extend it to energy and telecoms. Specifically, the government wanted to enable consumers to download their historical bank transaction data in a standardized format so they could use price comparison tools to see whether they could save by moving to another bank. By opening up this information that banks might otherwise have unique access to, the government hoped that customers would switch bank accounts more frequently, the way they do with credit cards, thus improving competition.[59]

The Midata initiative first rolled out in 2014 and became compulsory in 2015, and so far it seems to be neither a significant success nor failure. The only major company that offers the hoped-for price comparison service is GoCompare,[60] and their tool is not without criticism.[61] The U.K. Department for Business Innovation & Skills has not released data on whether bank account switching has increased.

However, Midata's standardization of how banks make transaction data available to users may lead to other benefits in the future. Previously, most U.K. banks permitted users to download their data in some form or another, but many made their data available only in a non-machine readable format, like PDF. Banks also varied widely in how far back in the transaction histo-

---

58.   *See* U.K. DEP'T FOR BUS. INNOVATION & SKILLS, PERSONAL DATA: REVIEW OF THE MIDATA VOLUNTARY PROGRAMME 3 (2014), https://assets.publishing.service.gov.uk /government/uploads/system/uploads/attachment_data/file/327845/bis-14-941-review-of-the-midata-voluntary-programme-revision-1.pdf.

59.   *See id.* at 3–4.

60.   *Money Comparison,* GOCOMPARE, https://www.gocompare.com/money (last visited Mar. 4, 2021). Numerous business leaders have praised the midata initiative (e.g., Martin Lewis from MoneySavingExpert, Alistair Crane from Grapple) but have yet to create their own comparison tools, so far as I can tell. *See* Press Release, U.K. DEP'T FOR BUS., INNOVATION & SKILLS ET AL.*, Businesses Get Creative with Consumer Data at the 'Midata' Innovation Lab Launch* (July 10, 2013), https://www.gov.uk/government/news/businesses-get-creative-with-consumer-data-at-the-midata-innovation-lab-launch.

61.   *Midata: Which? First Look*, WHICH? (Mar. 26, 2015), https://www.which.co.uk /news/2015/03/midata-which-first-look-399235. It also is unclear how much the tool is actually used, and since its initial release GoCompare has reduced the prominence of the midata tool on their website.

ry users could download data from (sometimes months, sometimes years). Midata also allowed the British government to set unified privacy standards for exported banking data.[62] Even if the data is underutilized now, standardization allows this data to potentially be used by new products in the future. The U.K. briefly ran the Midata Innovation Lab to explore potential uses for open data in banking and other industries.[63]

Sometimes, platforms can create a barrier to entry by having uniquely fast access to data.[64] Stucke and Grunes discuss this in terms of "nowcasting,"[65] the ability for a company to use the velocity of data they receive to identify trends before others. For example, Lyft's dynamic pricing works by increasing the price when it notices certain patterns of orders occurring in an area (e.g., people trickling out of a concert early before a massive rush).[66] The combination of timely data and broad historical data allows Lyft to more quickly and accurately adjust their prices to a customer's willingness to pay than a rival app with fewer users.[67] Twitter may use a similar process to identify trending topics faster than other social media sites, thus further perpetuating itself as the place to go for the most up-to-date conversation.[68]

One-off exports cannot address this kind of speed-based barrier to entry, but APIs can. As mentioned before, one-off exports take a long time to prepare all of a user's data, anywhere from minutes to months. APIs allow for smaller chunks of information to be accessed more quickly (often in a fraction of a second) and more frequently.

C. *Economies of Scale*

Economies of scale occur when a firm increases in efficiency as it increases in size. Usually, these gains in purchasing, managerial, and marketing power are eventually curbed by diseconomies of scale. Platforms that improve and grow from increased data collection, however, experience un-

---

62.   ATM transactions are unredacted, debit card transactions and fees redact card number data, and other transactions redact all descriptive data. Midata for Personal Current Accounts, MIDATA, http://www.pcamidata.co.uk [https://web.archive.org/web/20190306215048 /http://www.pcamidata.co.uk/].

63.   *See generally* MIDATA INNOVATION LAB, http://www.midatalab.org.uk (last visited Nov. 12, 2020).

64.   *See* MAURICE E. STUCKE & ALLEN P. GRUNES, BIG DATA AND COMPETITION POLICY 162–86 (2016); Rubinfeld & Gal, *supra* note 47, at 353.

65.   *See* STUCKE & GRUNES, *supra* note 64, at 8.

66.   *See* Davide Crapis & Chris Sholley, *Dynamic Pricing to Sustain Marketplace Balance*, MEDIUM (Nov. 10, 2020), https://eng.lyft.com/dynamic-pricing-to-sustain-marketplace-balance-1d23a8d1be90.

67.   *Id.*; *see also* Dawn Woodward, *Matching and Dynamic Pricing in Ride-Hailing Platforms*, MICROSOFT RSCH. (May 1, 2018), https://www.microsoft.com/en-us/research /video/matching-dynamic-pricing-ride-hailing-platforms.

68.   *See* Shahab Saquib & Rashid Ali, *Understanding Dynamics of Trending Topics in Twitter*, INTERNATIONAL CONFERENCE ON COMPUTING, COMMUNICATION AND AUTOMATION, May 5–6, 2017, at 98–103.

bounded economies of scale.[69] Stucke and Grunes describe how the dynamic plays out:

> [T]he more people actively or passively contribute data, the more the company can improve the quality of its product, the more attractive the product is to other users, the more data the company has to further improve its product, which becomes more attractive to prospective users.[70]

Growth and informational economies of scale are paramount virtues in the culture of Silicon Valley.[71] Many platforms only become profitable once they reach a certain scale or market share, so they depend on venture capital, often for years, to offset their high upfront costs.

Many scholars, namely those subscribing to the Chicago School view, argue that economies of scale do not themselves constitute barriers to entry.[72] Others argue that the drive to reach and maintain economies of scale incentivize certain barrier creating behaviors. Khan describes two such behaviors in the context of Amazon: predatory pricing and vertical integration.[73] Firms may even go so far as to take data from other platforms, as Google Maps did when it scraped images from Yelp to increase the number of photos it had for businesses.[74]

Data portability can only help competitors overcome information economies of scale if it is implemented in a way that can bring a lot of data and users away from the data sender to the data receiver. There is not much of a difference between APIs and one-off exports on this front. Though APIs help receivers save on storage costs and provide access to up-to-date, targeted information, the relative economies of scale between the receiver and the sender may not change as much because the data continues to exist on the incumbent platform. Even if users want to move away from the incumbent platform, informational economies of scale may mean that competitors' services are vastly inferior to the incumbent's until those users actually move, creating a coordinated action chicken-and-egg problem.

---

69.   *See* Robert Wilson, *Informational Economies of Scale*, 6 BELL J. ECON. 184, 184–95 (1975), https://www.jstor.org/stable/3003221.

70.   STUCKE & GRUNES, *supra* note 64, at 170.

71.   *See generally* PETER THIEL & BLAKE MASTERS, ZERO TO ONE: NOTES ON STARTUPS, OR HOW TO BUILD THE FUTURE, CROWN BUSINESS (2014)*;* BEN HOROWITZ, THE HARD THING ABOUT HARD THINGS: BUILDING A BUSINESS WHEN THERE ARE NO EASY ANSWERS, HARPER BUSINESS (2014).

72.   MARC A. EISNER, ANTITRUST AND THE TRIUMPH OF ECONOMICS: INSTITUTIONS, EXPERTISE, AND POLICY CHANGE 105 (1991).

73.   Khan, *supra* note 2, at 722–37. *But see* Daniel Crane, *The Tempting of Antitrust: Robert Bork and the Goals of Antitrust Policy*, 79 ANTITRUST L.J. 835, 852 (2014).

74.   *See* Natasha Tiku, *Yelp Claims Google Broke Promise to Antitrust Regulators*, WIRED (Sept. 12, 2017, 2:01 PM), https://www.wired.com/story/yelp-claims-google-broke-promise-to-antitrust-regulators.

Platforms have successfully used data portability to bootstrap their growth, but often at the expense of other normative values. LinkedIn adopted a legally dubious approach to data portability, colloquially known as "growth hacking," shortly after its founding in 2002.[75] Unlike Friendster, an open social network and the most popular at the time, LinkedIn was meant to connect professionals who knew each other in real life, allowing users to better trust recommendations made on the platform. This gated system assured that connections between users were meaningful, but it also meant less of them, which translated into a sparser network.[76] LinkedIn's revenue streams (job listings, subscriptions for improved communications, advertisements) all depended on the platform having many users and connections, so it needed to find a way to grow.[77]

To bootstrap its growth, LinkedIn ported contact lists from email clients, a novel strategy at the time.[78] It built a plugin for the Microsoft Outlook email client that mined a user's contacts and sent out invitations to people who had not already signed up for LinkedIn. In 2011, LinkedIn allowed users to upload their email contact lists from other clients with its "Add Connections" feature, and soon, up to 7% of users were uploading their address books, thus increasing the number of invitations sent to join LinkedIn by more than 30%.[79] LinkedIn gained access to even more contact information by acquiring Rapportive, another product that vacuumed up data from email clients.[80]

LinkedIn eventually sent out so many invitations that it ran into legal trouble. When someone who was invited to LinkedIn through the Add Connections feature did not respond, LinkedIn would follow up two more times, without explicit permission from the inviter.[81] If a person who was not a member of LinkedIn appeared in multiple uploaded contact lists, that could mean a huge volume of email. LinkedIn faced a lawsuit over this activity in 2015 and paid out $13 million in a settlement.[82] By then though, LinkedIn had already leveraged the scale of email to create its own scale. It had

---

75.   *See* Morgan Brown, *LinkedIn Growth Engine: The Never Ending Viral Loop*, GROWTHHACKERS (2014), https://growthhackers.com/growth-studies/linkedin.

76.   *See id.*

77.   *See* Ellen Lee, *LinkedIn's Startup Story: Connecting the Business World*, CNN MONEY (June 2, 2009, 10:24 AM), https://money.cnn.com/2009/06/02/smallbusiness /linkedin_startup_story.smb/index.htm [https://web.archive.org/web/20180327012620/https:/ /money.cnn.com/2009/06/02/smallbusiness/linkedin_startup_story.smb/index.htm].

78.   Brown, *supra* note 75.

79.   *Id.*

80.   *See* Anthony Ha, *Rapportive Announces Acquisition by LinkedIn, (Basically) Confirms $15M Price*, TECHCRUNCH (Feb. 22, 2012, 2:04 PM), https://techcrunch.com/2012/02 /22/rapportive-linkedin-acquisition.

81.   Jacob Kastreakes, *LinkedIn Agrees to Settle Unwanted Email Lawsuit*, VERGE (Oct. 2, 2015, 8:47 PM), https://www.theverge.com/2015/10/2/9444067/linkedin-email-lawsuit-settlement-add-connections.

82.   *Id.*

grown consistently, from 37 million members in 2009 to over 400 million members in 2016 when Microsoft acquired the company for $26.2 billion.[83]

The growth story of LinkedIn shows how portability, especially when done shrewdly and through backchannels, can improve competition at the expense of end users. The "growth hacking" approach raises serious questions about privacy, security, and spam. A portability regime should allow for competition to thrive without impinging the rights and control of users over their data.

D. *Network Effects*

Network effects occur when a product or service increases in value as more people use it. If one platform grows far larger than its competitors, that means network effects may entrench their advantage, creating a barrier to entry for others.[84] There are two types of positive network effects: *direct* and *indirect*.[85] Direct network effects occur when economic agents value a platform more when similar economic agents use it, like how telephone users gained value when others had telephones since it meant more people they could call. Similarly, Twitter benefits from positive network effects, since more active users on Twitter leads to more relevant and timely tweets. Indirect positive network effects occur when an economic agent benefits from a different kind of agent using the platform. For example, Lyft drivers benefit from more riders on the app and vice versa, as drivers can more quickly find passengers and passengers have shorter pick up times. A failure to capitalize on network effects leads to a platform's demise—a new ridesharing app will have a hard time attracting drivers without riders and vice versa; a rival to Twitter will have difficulty establishing itself as a dominant public forum if no one is there.

Prominent data economist David S. Evans outlines strategies in his catalyst framework for new platforms entering the market to build up their own network effects.[86] The value of multisided platforms, he argues, is that they bring together distinct economic agents, reducing transaction costs for them to find one another. The difficult part for platforms is finding the initial catalyst to get those economic agents onto the platform in the first place.

This challenge takes many forms. Sometimes, a platform needs to attract economic agents to one side of the market first. For example, many social media platforms match users with advertisers but only need users to ini-

83.   *Numbers of LinkedIn Members From 1st Quarter 2009 to 3rd Quarter 2016*, STATISTA (Oct. 27, 2016), https://www.statista.com/statistics/274050/quarterly-numbers-of-linkedin-members.

84.   *See* STUCKE & GRUNES, *supra* note 64, at 162.

85.   Michael Katz & Carl Shapiro, *Systems Competition and Network Effects*, 8 J. ECON. PERSPS. 93, 95–96 (1994).

86.   DAVID EVANS, PLATFORM ECONOMICS: ESSAYS ON MULTI-SIDED BUSINESSES 58–62 (2011).

tially get their business off the ground (so long as they have funding to run their service). In other cases, both sides of the market need to be catalyzed at once. A dating app that targets heterosexual individuals, for example, would need to have both straight men and straight women on the app to encourage the other to join. And some platforms need agents on one side to make a large, disproportionately risky initial investment. Video game systems, for example, must convince developers to make a large investment in their system by building games for it, often even before the system has reached the market and has any users.

Evans offers numerous proposals for addressing catalytic problems, and they all involve coordinating groups of economic agents to move onto the platform together.[87] One strategy he suggests is contingent contracts, which encourages one side to join after certain conditions have been met. Another approach is the "zig-zag strategy", in which the platform alternates onboarding chunks of users onto each side. A third Evans approach is the marquee strategy, in which the platform tries to attract a prestige user who will bring a following with them.[88]

Competitors could potentially use data portability as a part of any of these strategies. However, if individuals are only able to export their data one user at a time, it may be too difficult to organize enough users for the platform to reach catalyzation. Rather, groups of users would need to be able to move their data together in coordinated movements. As of today, no major platform offers a data portability tool that allows groups of users to migrate en masse to a new platform.[89]

One might argue that platforms could use one-off exports to address catalytic problems and network effects by allowing users to download their social graphs.[90] Facebook allows something similar in its one-off export tool Download Your Information—users can download a list of their friends which allows them to match with any friend who uploads their own list.[91]

---

87.   *Id.*
88.   Microsoft employed this strategy with their game streaming when they paid Ninja, a popular streamer of the most streamed game Fortnite, to move exclusively to their platform from Amazon's platform, Twitch. *See* Tramel Raggs, *Fortnite Star Ninja Leaves Twitch, to Stream on Microsoft's Mixer*, WASH. POST (Aug. 1, 2019, 6:18 PM), https://www.washingtonpost.com/sports/2019/08/01/fortnite-star-ninja-leaves-twitch-stream-microsofts-mixer.
89.   *See infra* Part IV.
90.   *See* Josh Constine, *Facebook Shouldn't Block You from Finding Friends on Competitors*, TECHCRUNCH (Apr. 13, 2018, 1:58 PM), https://techcrunch.com/2018/04/13/free-the-social-graph; *see generally* FACEBOOK PORTABILITY WHITE PAPER, *supra* note 35.
91.   A brief technical note: when users download a list of their friends, in a file called *friends.json*, they get their name and the timestamp of their friendship. This combined with the name of the user who has uploaded their friends list could more or less work to connect two users if they both uploaded their friends list, without giving any information about mutual friends who did not consent. This is called an edge export. *See generally* FACEBOOK PORTABILITY WHITE PAPER, *supra* note 35. However, the identifier is not strictly unique—if

Without access to external contact information though,[92] the exported social graph cannot fuel growth. It only becomes useful once a platform has a large number of users, at which point it will benefit from its own network effects anyways.

APIs address network effects better than one-off exports only to the extent that data senders are comfortable sending more data knowing that they can cut off access.[93] This data is limited in usefulness unless it also includes external contact data, which in the wake of Cambridge Analytica[94] and other privacy scandals happens less often. In general, though, APIs do not inherently make it easier for users to coordinate their movements or for platforms to overcome network effects.

III. ADDITIONAL CONSIDERATIONS FOR API PORTABILITY

The previous section mentioned multiple advantages that APIs have over one-off exports in helping new software services overcome barriers to competition. However, APIs come with additional factors regulators should consider when creating laws or guidelines around data portability. The main difference between one-off exports and APIs is the ongoing relationship between the data sender and receiver. One-off export regimes do not need to consider this relationship since the user acts as an intermediary between the two. But API regimes must consider this relationship carefully (especially because new products may come to depend on them) and regulators may have to step in to balance the power dynamics between them.

This section describes three additional factors specific to APIs and competition that regulators should consider. First, competitors may be hesitant to build a product that depends on incumbent APIs out of fear that the

---

two John Mills each became friends with two Li Weis at the exact same second, the identifier would not be unique. *See* NICHOLAS & WEINBERG, *supra* note 24.

92.   I looked at my own *friends.json* file and found that it listed an email address or phone number for 1.5% of friends in that list. Facebook does not document what data a user should expect to see in their DYI export, so I do not know why the contact information is listed for some friends but not others, but it may have to do with their privacy settings. *See* NICHOLAS & WEINBERG, *supra* note 24.

93.   *E.g.*, Facebook exhibits this behavior with data they share about a user's friends. In their one-off export tool (Download Your Information), users can only download information that includes each of their friends' names and the time they became friends. Their API (the Graph API) also gives a unique ID, but limits the data to friends who have also installed the application. *See Graph API Reference*, FACEBOOK FOR DEVS., https://developers. facebook.com/docs/graph-api/reference/v7.0/user/friends (last visited Feb. 16, 2021).

94.   *See* Matthew Rosenberg, Nicholas Confessore & Carole Cadwalladr, *How Trump Consultants Exploited the Facebook Data of Millions*, N.Y. TIMES (Mar. 17, 2018), https:/ /www.nytimes.com/2018/03/17/us/politics/cambridge-analytica-trump-campaign.html;   Ime Archibong, *New Facebook Platform Product Changes and Policy Updates*, FACEBOOK FOR DEVS. (Apr. 24, 2018), https://developers.facebook.com/blog/post/2018/04/24/new-facebook-platform-product-changes-policy-updates.

incumbent will change them or cut off access. Second, APIs may not lead to products that are sufficiently differentiated from the incumbent, and users either may not switch to competitors in the first place or may backslide to the incumbent. Third, APIs may better serve complementary platforms rather than analogous direct competitors.

A. *Hesitance to Depend on Incumbent APIs*

API portability potentially allows competitors to receive more frequent, up-to-date data from incumbents, but it also creates a complicated power dynamic between data senders and receivers for regulators to manage. From a technical standpoint, data senders have complete control over who can receive data and what data they can receive, and historically platforms have used this power to withhold data from threatening competitors. Until recently, Facebook explicitly disallowed applications that "replicate core functionality that Facebook already provides" in the platform policy for its social Graph API.[95] It used this justification to cut off access to multiple social media competitors, including Vine,[96] Yandex Wonder,[97] and Voxer.[98] Facebook did something similar to Prisma, a popular image filter app from 2016 that used a machine learning technique called "style transfer."[99] In November 2016, Prisma allowed users to add its filters on Facebook Live streams. Within the same month, Facebook revoked Prisma's ability to use the Facebook Live API and released its own style transfer feature.[100]

Other cases of Facebook restricting API access fall in the gray space between user protection and anti-competitiveness. Zynga used multiple Facebook APIs to grow a massive gaming platform on top of Facebook itself.

---

95.   Josh Constine, *Facebook Ends Platform Policy Banning Apps That Copy Its Features*, TECHCRUNCH (Dec. 4, 2018, 6:09 PM), https://techcrunch.com/2018/12/04/facebook-allows-competitors.

96.   According to emails released by the UK Parliament as part of the Six4Three case, this was personally approved by Mark Zuckerberg the day Vine launched. DIGITAL, CULTURE, MEDIA AND SPORT COMMITTEE, NOTE BY DAMIAN COLLINS MP, CHAIR OF THE DCMS COMMITTEE: SUMMARY OF KEY ISSUES FROM THE SIX4THREE FILES (UK), https://www.parliament.uk/documents/commons-committees/culture-media-and-sport/Note-by-Chair-and-selected-documents-ordered-from-Six4Three.pdf.

97.   *See* Josh Constine, *Facebook Blocks Yandex's New Social Search App from Accessing Its Data Just Three Hours After Launch*, TECHCRUNCH (Jan. 24, 2013, 4:49 PM), https://techcrunch.com/2013/01/24/facebook-blocks-yandex-wonder.

98.   *See* Josh Constine, *Facebook Is Cutting off Find Friends Data to "Competing" Apps That Don't Share Much Back, Starting with Voxer*, TECHCRUNCH (Jan. 18, 2013, 3:40 PM), https://techcrunch.com/2013/01/18/facebook-data-voxer.

99.   *See generally* Vlad Savov, *Prisma Will Make You Fall in Love with Photo Filters All over Again*, VERGE (July 19, 2016, 9:09 AM), https://www.theverge.com/2016/7/19/12222112/prisma-art-photo-app; Natasha Lomas, *Facebook Has Cut off Prisma's Live Video Access*, TECHCRUNCH (Nov. 30, 2016, 6:12 AM), https://techcrunch.com/2016/11/30/facebook-has-cut-off-prismas-live-video-access.

100.   Lomas, *supra* note 99.

The two companies had a symbiotic relationship, with Zynga accounting for 19 percent of Facebook's revenue[101] and Facebook accounting for 93 percent of Zynga's in 2011.[102] However, their relationship soured: Zynga was upset when Facebook started to take 30 percent of its revenue through its "Facebook Credits" system,[103] and Facebook was upset with Zynga for spamming its users with game invites. As Zuckerberg said in an interview:

> A lot of users like playing games, but a lot of users just hate games, and that made it a big challenge, because people who like playing games wanted to post updates about their farm or frontier or whatever to their stream. But people who don't care about games want no updates. So we did some rebalancing so that if you aren't a game player you're getting less updates.[104]

In 2012, Facebook ended its special relationship with Zynga and introduced API changes that hampered Zynga's business model.[105] Zynga was no longer able to have users send out mass game invites to their friends, and shortly thereafter users could not invite friends altogether. Around this time, Zynga's valuation dropped from $15 billion to $3 billion in six months.

The stories of Zynga and other competitors cut off from the Facebook API serve as a warning to startups and the venture capitalists who fund them—a platform that depends on a third-party API for growth may fail if the API changes. If regulators try to improve competition by enforcing API portability, they would have to figure out how to strike a balance between preventing incumbents from breaking competing platforms that depend on their APIs and letting those incumbents innovate and update their services. The ideal balance may change dramatically between different types of products, making scalable policy difficult.

---

101.   Tomio Geron, *Facebook's Dependence on Zynga Drops, Zynga's Revenue to Facebook Flat*, FORBES (July 31, 2012, 8:10 PM), https://www.forbes.com/sites/tomiogeron/2012/07/31/facebooks-dependence-on-zynga-drops-zyngas-revenue-to-facebook-flat.

102.   Zynga Inc., Quarterly Report (Form 10-Q) (Oct. 26, 2012).

103.   Zynga also threatened to create their own platform. Michael Arrington, *Zynga Gunning up (and Lawyering up) for War Against Facebook with Zynga Live*, TECHCRUNCH (May 7, 2010, 2:59 PM), https://techcrunch.com/2010/05/07/zynga-gunning-up-and-lawyering-up-for-war-against-facebook-with-zynga-live.

104*.   Exclusive: Discussing the Future of Facebook and the Facebook Ecosystem with CEO Mark Zuckerberg*, ADWEEK (June 22, 2010), https://www.adweek.com/performance-marketing/exclusive-discussing-the-future-of-facebook-and-the-facebook-ecosystem-with-ceo-mark-zuckerberg.

105.   Tomio Geron, *Zynga and Facebook Revamp Deal, Zynga Shares Drop*, FORBES (Nov. 29, 2012, 7:25 PM), https://www.forbes.com/sites/tomiogeron/2012/11/29/zynga-and-facebook-revamp-deal-zynga-shares-drop.

B. *Backslide and Undifferentiated Products*

If an incumbent offers an API or other means of interoperability, competitors that hope to build a similar product or service to that incumbent may have difficulty differentiating their offering, especially once price has reached an equilibrium.[106] If an incumbent and a competitor have nearly identical products, even low switching costs will not encourage users to change to the competitor. This occurred when the Federal Communications Commission (FCC) opened up competition between local exchange carriers (LECs, predominantly the seven Regional Bell Operating Companies, or "Baby Bells") and long-distance companies (e.g., AT&T, MCI, Sprint) in the Telecommunications Act of 1996.[107]

After the divestiture from AT&T in 1984, LECs were regulated as natural monopolies and long-distance companies were left to openly compete in other areas of telecommunications. When the FCC passed the Telecommunications Act of 1996, it wanted LECs and long-distance companies to compete in each other's markets. In order to give the long-distance companies a fighting chance against the LECs, the Act required LECs to provide them with "nondiscriminatory access to network elements on an unbundled basis"[108] (a data sharing access scheme similar to an open API) at wholesale price. Only once state authorities, the FCC, and the Justice Department all agreed that the LECs had adequately given access to their systems[109] would the LECs in turn be allowed to go into the long-distance market. This gave the long-distance companies a four-year head start to selling packages that bundled local and long-distance service.

Despite allowing for competition on more even footing, the 1996 Act did not create long term sustainable competition in telecommunications. LECs and long-distance companies offered essentially the same product, with similar costs and revenue, only adding the minor convenience of giving customers one bill instead of two. Very few customers actually switched their providers, and those who switched when telecoms started paying customers to do so, often switched back when other telecoms made the same offer.[110] Wireless eventually destroyed the profits of the long-distance indus-

---

106.   *See, e.g.*, Luuk de Klein & Lisha Zhou, *French, Australian Agency Chiefs Share Ideas on Privacy, Competition Interplay*, in ABA's 67TH ANTITRUST LAW SPRING MEETING CONFERENCE COVERAGE 18, 18 (PaRR ed., 2019), https://www.acuris.com/assets /PaRRABAreport2019_1.pdf.

107.   *See* PETER W. HUBER, MICHAEL K. KELLOGG & JOHN THORNE, FEDERAL TELECOMMUNICATIONS LAW § 5.11.2.2 (2d ed. 2011).

108.   47 U.S.C.S. § 251(c)(3) (LEXIS through Pub. L. No. 116–182).

109.   *See generally* 47 U.S.C.S. § 251 (LEXIS through Pub. L. No. 116–182).

110.   Interview with Randy Milch, former General Counsel and Head of Public Policy, Verizon Commc'ns, in N.Y.C., N.Y. (Nov. 15, 2019); *see generally* Jerry A. Hausman & Gregory J. Sikand, *Did Mandatory Unbundling Achieve Its Purpose? Empirical Evidence from Five Countries*, 1 J. COMPETITION L. & ECON. 173 (2005).

try, and the long-distance companies failed to diversify enough to save their business model. Most of them were cheaply acquired by LECs, further concentrating the telecommunications industry.[111]

The 1996 Act was a success in many ways, but the data portability aspect of it failed because customers were not willing to bear even a low switching cost between offerings that were so similar. Customers that were willing to switch—for example, because a competitor offered sign up bonus—could be induced to switch back if the incumbent offered the same bonus. The competitors, here the long-distance companies, were limited in how much they could differentiate their products so much because they were all built on top of the same local exchange connections. Competitors that tie their products too closely to incumbent APIs may face a similar issue.

C. *Opportunities for Creative Destruction*

Despite its issues, API portability still has the potential to improve the competitive landscape of the tech sector. To better understand how, it is useful to look at two different notions of competition: *neoclassical* and *Schumpeterian*. Neoclassical economics defines competition in terms of the desired outcomes of improved product and lower price.[112] This approach to competition is easy to model and promotes creating directly competing products over wholesale innovation. Early 20th century economist Joseph Schumpeter on the other hand saw competition as a process of "creative destruction," whereby innovative newcomers disrupt and dislodge incumbents.[113] It is a process of explosive, unpredictable change that leads to a larger paradigm shift towards complementary products rather than directly competing ones.

API portability may not work well for encouraging neoclassical competition, as shown by the example of the Telecommunications Act of 1996, but it can encourage Schumpeterian competition. Innovative, complementary products are not as easily harmed by backslide and switching costs. And even if data senders eventually copy the product themselves, data receivers retain a first-mover advantage. This dynamic played out in dating apps—Tinder used the Facebook API to make it easier for users to build their profiles with ported Facebook data. When Facebook released its own dating feature later on (again, it took longer to build a complementary product),

---

111.   *See generally* Hausman & Sikand, *supra* note 110.

112.   *See generally* Herbert Hovenkamp, *United States Competition Policy in Crisis: 1890-1955*, 94 MINN. L. REV. 311 (2009).

113.   Joseph A. Schumpeter, CAPITALISM, SOCIALISM AND DEMOCRACY 81–86 (Taylor & Francis ed., 2003) (1942).

Tinder built up enough of a brand name and network effect that it wasn't a threat.[114]

The potential for Schumpeterian style competition through APIs can be seen with ported banking data in the financial tech sector.[115] Companies like Mint, Yodlee, and Quicken leverage banking data, often from multiple sources, to give users a comprehensive picture of their finances. Another fintech company, Cushion, reads a user's bank transaction history and negotiates lower fees on her behalf. Plaid, acquired by Visa in 2020 for $5.3 billion, created APIs that allow third party developers to interface with banks to further allow this industry to grow.[116]

All of this growth occurred without banks and tech companies cooperating. Open banking is an example of interoperability without explicitly permitted APIs—fintech companies often use unsanctioned logins and screen scraping to collect financial data[117] while banks in turn improve their anti-scraping technology.[118] At the same time, banks sue tech companies over "exceeding authorized access" under the Computer Fraud and Abuse Act[119] and protection of "nonpublic personal information" under the Gramm-Leach-Bliley Act.[120] Tech companies meanwhile insist that the Dodd-Frank Act protects their use of consumer data.[121] Given the large technical and financial burden of Mint and Plaid-style jerry-rigged interoperability, only a fraction of the innovation that would be possible with officially sanctioned APIs is represented.

Galvanizing Schumpeterian and neoclassical competition each require different approaches,[122] and both have their positives and negatives. Neoclassical competition is easier to regulate but is unlikely to bring about

---

114.   Kurt Wagner, *Match Group Says Facebook's New Dating Feature Will Have 'No Negative Impact on Tinder'*, RECODE (May 8, 2018, 4:55 PM), https://www.vox.com/2018/5 /8/17332644/match-group-facebook-earnings-tinder-dating-competition. CEO of IAC Joey Levin said in response to the release of Facebook Dating, "Come on in. The water's warm. Their product could be great for US/Russia relationships."

115.   *See generally* Nizan Geslevich Packin, *Show Me the (Data About the) Money!*, 5 UTAH L. REV. 1277 (2020).

116.   *See id.* at 1287; *Visa to Acquire Plaid*, VISA (Jan. 13, 2020), https://usa.visa.com /about-visa/newsroom/press-releases.releaseId.16856.html.

117.   For example, Plaid using link injections to get bank credentials. *See* Ben Ellis (benissimo), *Privacy/Security Concerns #68*, GITHUB (Feb. 11, 2016), https://web.archive.org /web/20190415103059/https://github.com/plaid/link/issues/68; Whockey, Comment to *Plaid Deletes GitHub Issue Exposing Imitation of Bank Login UIs*, HACKER NEWS (June 8, 2019), https://news.ycombinator.com/item?id=20133806.

118.   *See* Nathan DiCamillo, *In Data Dispute with Capital One, Plaid Stands Alone*, AM. BANKER (July 17, 2018, 3:52 PM), https://www.americanbanker.com/news/in-data-dispute-with-capital-one-plaid-stands-alone.

119.   18 U.S.C.S. § 1030(a)(1) (LEXIS through Pub. L. No. 116–182).

120.   15 U.S.C.S. §§ 6801–6809 (LEXIS through Pub. L. No. 116–182).

121.   *See generally* 12 U.S.C.S § 5481 (LEXIS through Pub. L. No. 116–182).

122.   *See generally* Tim Wu, *Taking Innovation Seriously: Antitrust Enforcement if Innovation Mattered Most*, 78 ANTITRUST L.J. 313 (2012).

widescale innovation. Schumpeterian competition creates significant innovative value for firms and consumers when it is successful, but it is difficult to predict. In terms of portability, it is hard for policymakers to know whether or not certain datasets have innovative potential and should therefore be made portable. Schumpeterian competition through API portability therefore must be designed on an industry by industry basis. I will set aside further interrogation of this problem for a future paper. For now, I will turn to a new, more scalable approach to data portability that is designed to maximally encourage neoclassical competition.

IV. COLLECTIVE PORTABILITY—A NEW APPROACH

The previous two sections reviewed approaches to data portability that focus on providing individual users access to their data. This reflects the ways current laws regulate portability and platforms implement portability. The individual approach works to alleviate switching costs and some unique data access barriers but fails to address network effects and informational economies of scale. Previous analysis of existing theories and examples of portability suggests these barriers can be better addressed by groups coordinating to move their data all at once.

To this end, I propose a new approach to portability called *collective portability*: the ability for users to easily transfer their data to another platform in its original unit and alongside other users they used it with. More specifically, if a group of users shares access to co-created data on a specific platform and they want to move that data to a new platform, the original platform should facilitate the group in coordinating the transfer (and potential deletion) of all their data at once. Like API portability, collective portability requires the data sender and receiver to be in direct communication with one another. However, since the relationship does not continue after the transfer is complete, a power imbalance between senders and receivers is not a concern.

Collective portability works better than one-off exports at lowering competitive barriers to entry. Not only does it better address network effects and informational economies of scale, but it also facilitates data transfer that otherwise could not occur without raising privacy concerns. One-off exports only allow users to transfer data that is clearly their own. If data is shared by multiple people, it may not end up in any of their exports. Collective portability includes this data in the transfer, making it more than the sum of every individual user's exports without infringing upon user privacy.

The shortcomings that collective portability addresses can be seen writ large in the data Facebook makes available in its Download Your Information tool. DYI allows users to download a computer-readable archive of all information they have ever entered into Facebook that it still has stored. The information includes posts, events attended, comments, advertisements

clicked on, and photos uploaded. It does not include data provided by other users, like tagged photos or likes from friends.[123]

I downloaded my personal DYI archive and below is an anonymized example of the data representing my comment on a photo that my friend (here named "Fannee Doolee") posted to a group to which we both belong (Figure 1):[124]

```
{
  "timestamp": 1597855333,
  "data": [{
    "comment": {
      "timestamp": 1597855333,
      "comment": "Wow this vacation looks awesome!",
      "author": "Gabriel Nicholas",
"group": "Little Wanderers: Traveling with Kids"
    }
  }],
  "title": "Gabriel Nicholas commented on Fannee Doolee's photo."
}
```

FIGURE 1

The data I received consists of the text of the comment, the time I posted it,[125] the type of item I commented on (a photo), the name of the group I posted to (Little Wanderers), and the name of the person who owns it (Fannee Doolee). I do not have access to the photo itself, or any unique identifi-

---

123. The information also does not include any inferences Facebook may have made about a user. For example, if Facebook has determined that a user supports gun owner rights and uses that information to show ads for guns, DYI does not give a user who downloads their information access to that insight. *See* FACEBOOK PORTABILITY WHITE PAPER, *supra* note 35, at 13–16.

124. For this analysis, I downloaded my own Facebook information, which goes from 2006 to 2019. I downloaded my data in the machine-readable JSON format as opposed to the more human-readable HTML format, since this is likely what actual competitors would use to integrate the data. My data does not represent the entire range of possible data to download because I have not used every feature on Facebook. For example, I have never used Facebook's "check in" feature, so my data download has an empty folder where that data would be. I also may have had certain privacy settings enabled that prevented certain types of information from being downloaded. People who I have interacted with on Facebook may also have these settings. All of this is speculative because Facebook provides no documentation about what data structures a user should expect from their DYI export. I reached out to Facebook to get basic information on the rules underlying DYI or even the basic structure of data I could expect, but after six months of back and forth, they did not provide any, nor did they give any information on how different privacy settings might affect the export. I therefore had to write a script to reverse engineer the structure myself. Any data snippets here are anonymized versions of those results, though I have kept my own name for the sake of clarity. *See generally* Gabriel Nicholas (gajeam), *Portability-Project*, GITHUB (Sept. 13, 2019), https://github.com/gajeam/Portability-Project [hereinafter *Portability Project GitHub*].

125. Note that though there are two timestamps, they are identical. *See generally id.*

ers for the group, person, or photo.[126] If Fannee Doolee downloaded all of her photos, she would similarly not have access to my comment. Therefore, even if Fannee Doolee and I both uploaded our data to the same competing photo-posting-and-commenting platform, it would not have a way to con-nect my comment to her photo.

Facebook may not have made this system design choice solely to be an-ti-competitive—it also addresses legitimate privacy concerns. If Fannee Doolee does not want to upload her photo to the platform that I upload my data to, she would not expect me to be able to bring it there without her permission.[127] And if a number of people who commented on that photo move to the new platform, she may not want the conversation around her photo moved somewhere without her permission. However, these privacy protections directly reduce the ability of competitors to recreate features of Facebook. DYI makes similar tradeoffs with other types of data, such as events, statuses, and likes.[128]

Collective portability would avoid this tradeoff between privacy and competitive utility, at least as it pertains to Facebook groups. If Little Wan-derers chose to move to another platform and Fannee Doolee and I both consented, under collective portability the new platform would be able to show the relationship between her photo and my comment, reducing switch-ing costs for us and increasing network effects for the new platform.[129]

Designing a regime for collective portability is not as straightforward as it is for one-off exports, even in this relatively simple Facebook example. Below, I outline five questions that platforms and policymakers would have to consider in order to implement and regulate collective portability, and I apply them to different hypothetical applications. I then look at the ways in which collective portability addresses the barriers to entry discussed in Part II. Finally, I point out new challenges collective portability raises.

A. *Five Questions for a Collective Data Portability Regime*

With any portability regime, platforms must make design decisions about which data they make available to users and how they make it availa-

---

126.   If I were friends with multiple people named "Fannee Doolee" or in multiple groups called "Little Wanderers", this would be a problem. *See generally id.*

127.   This could be considered an "inappropriate information flow". *See generally* HELEN NISSENBAUM, PRIVACY IN CONTEXT: TECHNOLOGY, POLICY, AND THE INTEGRITY OF SOCIAL LIFE (2009).

128.   Notably, it does not occur in data about private messages—users get both sides of the conversation and all reactions to messages, meaning that a competitor could in theory use this information to build a new messaging platform. *See generally Portability Project GitHub*, *supra* note 124.

129.   The *profile_information.json* file included among other things, my birthday, email address, current city, hometown, family members, education and work experiences, and pages I had liked ten years ago. *See generally id.*

ble. Collective portability raises additional design questions for a platform because it involves multiple users, with many potential relationships between each other and their data. Five such questions are discussed below:

1. *Who is in in the collective?* For many platforms, groups will have clear boundaries—it usually is clear who is and is not on a group messaging thread, for example. Some platforms may have more informal boundaries around their groups and communities, or no boundaries at all. For example, nothing strictly delineates which users are and are not part of "legal Twitter."[130] Twitter, therefore, may not be well-suited for collective portability.

2. *Which data is part of the collective?* Again, this is an easy question for smaller, simpler groups like message threads, but more difficult for larger, more fluid groups. For example, a community platform like Reddit would have to decide whether or not it allowed groups to port data cross-posted from another community.[131]

3. *How does the collective come to a decision to port their data?* Collective portability requires the data sender to implement a governance mechanism to allow users in the group to decide to move to another platform. There is no one-size-fits-all design for this mechanism, especially for larger groups. If the data sender allows a group to vote on porting their data elsewhere, how many users have to vote for the transfer to occur? How are inactive group members handled? What happens to data from users who voted against the move?

4. *How do special user roles affect the transfer decision process?* Some platforms provide certain users in a group more privileges than others. This makes some portability decisions trivial— if an employer wants to switch their team communication tool, she can make a unilateral decision to transfer group data. But complications may arise in non-professional communities, where administrators and moderators have less unilateral power.

---

130.   *But see* Sarah Mui, *Best Legal Twitter Accounts of 2018*, A.B.A. J. (Dec. 1, 2018, 12:15 AM), http://www.abajournal.com/magazine/article/best_law_twitter_2018.

131.   u/EuroGeologist, *How to Crosspost on Reddit in 2018?*, REDDIT (May 18, 2018, 4:54 AM), https://www.reddit.com/r/help/comments/8kbptp/how_to_crosspost_on_reddit_in_2018.

5. *How does the migration work?* Once decisions have been reached about which users and which data ought to be transferred, it is up to the data sender to pass the data along to the receiver. Data receivers may have to make tough technical and interface designs to handle groups with missing or incomplete data due to some users opting out of the migration. Similarly, the data receiver's schema for representing data may not map perfectly onto the sender's, especially if the platforms do not share all the same features. Receivers need to decide whether to discard this data or handle it another way.

B. *Hypothetical Collective Portability Regimes*

In this section, I apply the five questions outlined above to hypothetical collective portability regimes for Spotify, iMessage, and Facebook. These hypotheticals increase in complexity and are meant to highlight the competitive advantages and the difficult-but-manageable challenges of implementing collective portability.

1. Spotify

Imagine I use Spotify to stream music but I want to switch to YouTube Music because they offer a similar product at a lower cost. Spotify has years' worth of data on my playlists and albums saved to my library. They also have my listening history, which means they can give me high quality recommendations for songs I might like.

I am the only user in this collective, since I largely interact with the Spotify by myself, and therefore I can make unilateral decisions about my data. The data in this collective includes my playlists, music library, and listening history. Even if friends follow some of my playlists, the boundaries between my data and theirs are clear. Network effects, at least in terms of interacting within small groups, are less of a concern because the primary value of Spotify comes not from the network of other users but interaction with the data itself (i.e. listening to music as opposed to interacting with friends).

If Spotify were to adopt a collective portability regime, it would facilitate me sending all three of these data sources to a competitor, like YouTube Music, and in turn, allow me to cancel, pause, or delete my Spotify account. On the data receiver's end, YouTube Music would have to map Spotify's representation of songs onto its own, something it would only implement if it believed the effort would actually grow its user base.

2. iMessage

Imagine I am in an iMessage texting thread with four close friends. We all agree that we want to move our thread to WhatsApp because I am about to switch from an iPhone to an Android and WhatsApp offers better inter-operating system messaging.

The people in this collective are me and my four friends, and the data is our thread, with every message, image, and reaction sent. With so few people in the thread, iMessage may require us to unanimously consent to transfer our data. The voting mechanism could be built into iMessage itself and users would be able to pass along their WhatsApp credentials to iMessage to allow it to make the transfer. Some information that does not translate between the two systems, such as message reactions (iMessage has them, WhatsApp does not) gets lost.

3. Meme Groups

Imagine I am in a ten-thousand person Facebook group called "Anti-trust Memes for Neo-Brandeisian Teens" where users post dozens of irreverent image macros about competition in the tech sector every day. The admins of the group fear that Facebook is preventing their content from showing up on people's newsfeeds, and suspiciously, new users seem unable to join. They want to move the group to a new platform MemeOn!, built specifically for anti-authoritarian meme groups.

The people in this collective are the group members, and the data is the posts in the group, including images and comments. The governance mechanism to collectively port this data may be a direct or liquid democracy.[132] Facebook would have to decide how to handle data from users who did not vote or voted against the move, though this issue would be easier for groups started after Facebook begins collective portability since they could set portability options from the beginning. Administrators may have a special status, potentially deciding when to call a vote or choosing how many people would need to approve the vote for the group to move. Migration would raise a number of technical and user interface design questions, though likely every migrating user would need to connect to their MemeOn! account (this could be tied in with voting). Old data could be held on Facebook for archival purposes or the group could split, with some staying on Facebook and others going to MemeOn!.

---

132.   *See generally* Steve Hardt & Lia C. R. Lopes, *Google Votes: A Liquid Democracy Experiment on a Corporate Social Network*, TECH. DISCLOSURE COMMONS (June 5, 2015), https://www.tdcommons.org/dpubs_series/79.

C. *Collective Portability and Barriers to Entry*

In this section, I compare how well collective portability alleviates each of the previously established four barriers to entry compared to the current approaches to portability, one-off exports and APIs.

1. Switching Costs

Collective portability reduces switching costs by making it so that users do not have to leave people they know from the old platform behind when they switch to a new one. This would be particularly helpful for nascent social networks, where pioneering users often abandon platforms, even ones experiencing viral growth, because they are essentially alone. [133] Collective portability also lowers switching costs for groups of users with shared data who under other portability regimes would have to leave shared data behind. However, if the governance or transfer mechanisms are poorly designed, users may not take advantage of collective portability and the benefits may not be realized.

2. Unique Data Access

Collective portability opens up unique data sources that competitors can use as the foundation for building new platforms. By moving users en masse and allowing them to bring their data, a new entrant can target a specific large group and create a product directly aimed at its needs in ways that a large multi-use platform cannot. For example, Facebook groups have been used by political activists to organize and patient groups to support one another, but the platform fails to meet either of their privacy needs. [134] Both groups started on Facebook to reach the largest number of people, but now, platforms may exist that better meet their needs, especially if growth is no longer their main priority. Some groups may be so large that competitors would even build new products specifically tailored to them if only they could access the data. Collective portability is thus necessary to give groups the same kind of choice afforded to individuals. [135]

---

133.   For example, Wikitribune, Ello, and Vero are all social media platforms that surged in popularity in response to different Facebook scandals. They all had waitlists when they first launched because they could not handle the scale of user interest, but all failed to maintain that level of growth, potentially because too few people were on them.

134.   *See generally* Madelyn Rose Sanfilippo & Katherine J. Strandburg, *Privacy Governing Knowledge in Public Facebook Groups for Political Activism*, INFO., COMMC'N & SOC'Y (Sept. 26, 2019), https://www.tandfonline.com/doi/full/10.1080/1369118X.2019. 1668458; *United States v. Facebook, Inc.*, No. 1:19-CV-2184 (D.D.C. July 24, 2019) (stipulated order), https://www.ftc.gov/system/files/documents/cases/182_3109_facebook_ order_filed_7-24-19.pdf.

135.   Collective portability cannot address the unique speed of data access, aka "nowcasting", because there is only a one-time connection between the data sender and receiver,

3. Economies of Scale

Collective portability better addresses economies of scale than other approaches to data portability because it allows for more data and more users to move over at once. This helps competing platforms reach a similar scale to incumbents more quickly.[136] It also allows for scale without threatening privacy, as LinkedIn's contact scraping did, or compromising utility to competitors, as Facebook's DYI tool does. Additionally, when a collective agrees to move their data together, less data falls between the cracks as it would with individual users moving with their own one-off exports.

4. Network Effects

Collective portability aligns closely with the catalytic strategies Evans recommends for overcoming network effects.[137] According to his framework, platforms create value by bringing distinct economic agents together and reducing the transaction costs of them finding one another, and new entrants to the market can only get the economic agents they need to kickstart growth by onboarding large chunks of users at once. Collective portability does exactly that. And for some platforms, such as messaging platforms, a few strong network connections may create network effects as powerful as many weaker connections.

D. *Challenges to Collective Portability*

Collective portability has some distinct advantages over current one-user-at-a-time portability approaches like one-off exports and APIs. However, regulation and implementation of collective portability, beyond being difficult, threatens to exacerbate other problems faced by the tech industry. Platforms may also resist collective portability for reasons that relate to competition and otherwise. Some of the obstacles collective portability may face are discussed below.[138]

---

not an ongoing connection where timely information gets sent back and forth. *See* STUCKE & GRUNES, *supra* note 64, at 8.

136.   It may also not take having the same amount of data to compete. There are eventually diminishing returns on the amount of data and the features of the algorithm itself superseded. *See generally* Xinran He et al., *Practical Lessons from Predicting Clicks on Ads at Facebook*, FACEBOOK RSCH. (Aug. 24, 2014), https://research.fb.com/publications/practical-lessons-from-predicting-clicks-on-ads-at-facebook.

137.   *See generally* EVANS, *supra* note 86.

138.   *See* Mike Masnick, *Protocols, Not Platforms: A Technological Approach to Free Speech*, KNIGHT FIRST AMENDMENT INST., COLUM. UNIV. (Aug. 21, 2019), https://knightcolumbia.org/content/protocols-not-platforms-a-technological-approach-to-free-speech. These problems are the same as the ones in Masnick's approach to platform free speech and competition. The two proposals are very different, but both involve making it easier to move data around.

1. Implementation Complexity

Platforms may have a difficult time implementing collective portability regimes, which in turn may lead to complex designs that users do not readily adopt. Some platforms may have difficulty determining which users or what data constitute a collective. Users may have different gradations of control over certain data (e.g., a Google Doc can have owners, editors, commenters, and viewers) and ownership is not always exactly clear (e.g., if a Twitter user quotes another user's public tweet, is that tweet now a part of his data?).[139]

Governance mechanisms can also be difficult to design, especially for large communities in which the choice to change platforms will most likely not be unanimous. Additional normative and technical questions arise for groups and platforms that were built before implementing collective portability. Only platforms that build collective portability in from the beginning can set rules at the time users sign up for a group or platform. If the rules determining which data gets sent are too complicated or not transparent enough, new market entrants will be hesitant to invest in integrating with it, leaving users with nowhere to bring their data.

2. Incumbent Resistance

Currently dominant platforms are likely uninterested in implementing collective portability since they may see it as sabotaging their own businesses. Collective portability is in some sense a zero-sum game with platforms vying for the same users, data, and attention.[140] In this respect, large platforms have more to lose, particularly when entire groups can move at once, meaning they may actively resist collective portability. Platforms may also not want to take responsibility for answering the hairy questions about how to send data, especially if the process upsets some users or requires significant investment in resources Platform resistance would thereby put the onus on regulators to design competition-inducing collective portability standards, and they might not have the technical knowhow to do so effectively.

Collective portability may more broadly run counter to the way that large technology firms conceptualize their platforms and design policy for their users. For example, Facebook sets rules for its users by following signals from individual actors (users and organizations).[141] Collective portabil-

---

139.   Twitter says, yes. When I downloaded my Twitter data, I got other people's retweets, including their pictures.

140.   It is too speculative to say that collective portability would increase the number of users on the Internet or the amount of time they spend on multi-sided platforms.

141.   *See* Matthias C. Kettemann & Wolfgang Schulz, *Setting Rules for 2.7 Billion. A (First) Look into Facebook's Norm-Making Systems: Results of a Pilot Study* (Working Pa-

ity would require a paradigm shift in Facebook's approach to public policy to address the needs of pluralities rather than individuals.

3. Content Moderation and Filter Bubbles

If users can easily organize and move their data from one platform to another en masse, de-platforming could become more difficult. It could make it easier for hate speech and other normatively undesirable behavior to regrow on a new platform when stomped out on another. This dynamic has played out on a smaller scale between Reddit, a news aggregator and social network, and Voat, an almost identical service. In 2015, Reddit closed down five of its communities (or "subreddits") for harassment, including r/FatPeopleHate with 150,000 members.[142] A controversy over free speech on the platform broke out and some moved to Voat, then run by a two person team, where they restarted some of the closed communities.[143] Voat's FatPeopleHate group at one point had 18,700 members.[144] A similar story played out with the far right, pro-Trump r/The_Donald subreddit, which was banned in June 2020 for abuse and hate speech.[145] The community had ample warning before the ban and had largely abandoned the subreddit months earlier[146] in favor of a homegrown Reddit clone called TheDonald.win, which is even more popular than Voat.[147]

If Reddit had collective portability, it could have made the migration of r/FatePeopleHate or r/The_Donald even easier, potentially allowing them to migrate more users and keep the historical content they were banned for. Furthermore, if Reddit knew that hateful communities could easily reconstitute on new platforms, it may have been less incentivized to remove these communities in the first place under the logic of "if we don't host them

---

pers of the Hans-Bredow-Institut, Work in Progress # 1), https://leibniz-hbi.de/uploads/media/default/cms/media/1soch5s_AP_WiP001InsideFacebook.pdf.

142.   *See* Emma Woollacott, *Users Flock to Voat as Reddit Shuts Harassing Groups*, FORBES (June 11, 2015, 7:51 AM), https://www.forbes.com/sites/emmawoollacott/2015/06/11/users-flock-to-voat-as-reddit-shuts-harassing-groups.

143.   *See* Adi Robertson, *Welcome to Voat: Reddit Killer, Troll Haven, and the Strange Face of Internet Free Speech*, VERGE (July 10, 2015, 9:11 AM), https://www.theverge.com/2015/7/10/8924415/voat-reddit-competitor-free-speech.

144.   *Id.*

145.   *See* Bobby Alynn, *Reddit Bans The_Donald, Forum of Nearly 800,000 Trump Fans, over Abusive Posts*, NPR (June 29, 2020, 5:10 PM), https://www.npr.org/2020/06/29/884819923/reddit-bans-the_donald-forum-of-nearly-800-000-trump-fans-over-abusive-posts.

146.   Reddit's usage statistics show the page had on average only 7,780 daily active users at the time. Elizabeth Culliford & Katie Paul, *Reddit Bans 'The_Donald' Forum Amid Broad Social Media Crackdown*, REUTERS (June 29, 2020, 1:50 PM), https://www.reuters.com/article/us-reddit-trump/reddit-bans-thedonald-forum-amid-broad-social-media-crackdown-idUSKBN2402K7.

147.   *See* Sarah Emerson, *Months Before Reddit Purge, The_Donald Users Created a New Home*, ONEZERO (July 2, 2020), https://onezero.medium.com/months-before-reddit-purge-the-donald-users-created-a-new-home-a732f79e4f04.

someone else will." Frictionless ideology-based migration could also in-crease filter bubbles[148] by making it easier for groups with shared ideologies to flock to the same platform and isolate themselves from other ideas.

The capacity to deplatform remains normatively ambivalent. Forums like 8chan, a site linked to multiple mass shootings known for harboring white supremacism and anti-Semitism, should obviously be able to be de-platformed, as should safe harbors for child pornography and terrorist ac-tivity. However, deplatforming raises free speech concerns,[149] such as when Alex Jones was banned from YouTube and Facebook, seriously diminishing traffic to his show "Infowars."[150] The extent to which collective portability allows for certain kinds of deplatforming depends on specific choices in im-plementation.

CONCLUSION

Data portability as it exists today has yet to demonstrate that it can im-prove competition in the tech sector. The GDPR has required platforms to allow users to download their personal data since 2018, yet no meaningful competitors have emerged from this effort. Successful predecessors to plat-form data portability, such as mobile number portability, only alleviated switching costs; new multi-sided platforms hoping to enter the market face additional barriers. Since data on an incumbent platform is mostly useful in the context of other users, one person porting their data in isolation is often insufficient to bootstrap competitors.

Collective portability addresses market barriers to entry beyond switch-ing costs. It allows users to coordinate to transfer their data with others and maintain the context it was used in, assuring competitors that they can use it for the same purpose. This may be an appealing tool for regulators hoping to improve platform competition, even if it comes with its own challenges around governance, content moderation, and industry buy-in.

What might good collective portability regulation look like? First, it would be entirely separate from data portability regulation aimed at giving users better control over their data. Singapore set a precedent for this in forthcoming amendments to its Personal Data Protection Act, which sepa-rate out the competition and data control aims of data portability into two

---

148.   *See generally* ELI PARISER, THE FILTER BUBBLE: WHAT THE INTERNET IS HIDING FROM YOU (2011).

149.   *See* Jonah Goldberg, *Deplatform on Social Media and Free Speech*, NAT'L REV. (May 13, 2019, 5:19 PM), https://www.nationalreview.com/corner/deplatforming-on-social-media-and-free-speech.

150.   *See* Jack Nicas, *Alex Jones Said Bans Would Strengthen Him. He Was Wrong.*, N.Y. TIMES (Sept. 4, 2018), https://www.nytimes.com/2018/09/04/technology/alex-jones-infowars-bans-traffic.html.

separate rights, the Data Portability Obligation[151] and the Access Obligation.[152] Second, it would differentiate between small, new market entrants and large incumbents. This prevents users and data from further concentrating on large platforms with strong network effects while also relieving under-resourced startups from excessive technical burden. Third, the regulation would create explicit carve outs for when portable user data overlaps. The GDPR states that if personal data overlaps between two or more data subjects, privacy and deletion rights supersede portability rights, potentially leaving none of the data subjects able to port it. A collective portability version of this policy would provide that if the data subjects were part of some closed off group, the platform should make it possible for all relevant data subjects to consent to a transfer or export.

---

151.   *See* SINGAPORE PERSONAL DATA PROTECTION COMMISSION, RESPONSE TO FEEDBACK ON THE PUBLIC CONSULTATION ON PROPOSED DATA PORTABILITY AND DATA INNOVATION PROVISIONS 5 (2020), https://www.pdpc.gov.sg/-/media/Files/PDPC/PDF-Files /Legislation-and-Guidelines/Response-to-Feedback-for-3rd-Public-Consultation-on-Data-Portability-Innovation-200120.pdf. The Data Portability Obligation is intended "to facilitate movement of consumer data from one service provider to another, so that consumers are better empowered to try out or move to new or competing service offerings." It requires data controllers to transmit data in a machine-readable format directly to another controller if so requested.

152.   *Id.* The Access Obligation "is intended to allow individuals to access and verify their personal data in an organisation's possession or under its control, and how their personal data has been used by the organisation."