

UNPACKING DATA PORTABILITY



BY CHRISTOPHER S. YOO¹



¹ John H. Chestnut Professor of Law, Communication, and Computer & Information Science and Founding Director of the Center for Technology, Innovation and Competition, University of Pennsylvania.

CPI ANTITRUST CHRONICLE NOVEMBER 2020

CPI Talks...

...with *Jon Leibowitz*



Data to Go: The FTC's Workshop on Data Portability

By *Guilherme Roschke & Andrea Zach*



Data Portability

By *Daniel L. Rubinfeld*



Using the Portability and Other Required Transfers Impact Assessment ("PORT-IA") in Antitrust Law

By *Peter Swire & John Snyder*



Unpacking Data Portability

By *Christopher S. Yoo*



The Opportunities and Limits of Data Portability for Stimulating Competition and Innovation

By *Inge Graef*



Online Search Competition and the Risk of Unintended Consequences of Data Access

By *Jordi Casanova*



The Impact of Data Portability on Platform Competition

By *Emanuele Giovannetti & Paolo Siciliani*



Data Portability Rights: Limits, Opportunities, and the Need for Going Beyond the Portability of Personal Data

By *Daniel Gill & Wolfgang Kerber*



Data Access and Portability and EU Competition Law

By *Björn Lundqvist*



Visit www.competitionpolicyinternational.com for access to these articles and more!

CPI Antitrust Chronicle November 2020

www.competitionpolicyinternational.com
Competition Policy International, Inc. 2020[©] Copying, reprinting, or distributing this article is forbidden by anyone other than the publisher or author.

I. INTRODUCTION

Data portability has become a hot topic in competition law. Now well established as a matter of privacy law by the enactment of the California Consumer Protection Act ("CCPA") and the European Union's General Data Protection Directive ("GDPR"), legislators and enforcement officials around the world have shown increasing interest in data portability as a competition law remedy. It was endorsed by recent high-profile reports issued by expert panels convened by the European Commission and the UK in 2019. The 2019 report released by the Australian Competition and Consumer Commission ("ACCC") was more circumspect, concluding that data portability was unlikely to provide any short-term benefits to competition in digital platform markets. At the same time, the ACCC promised to revisit the issue when considering how to apply the 2017 consumer data right to other sectors.

The topic began to attract interest in the U.S. this fall. For example, the Federal Trade Commission conducted a workshop on the topic on September 20. In addition, the October 6 majority staff report that capped off the U.S. House Judiciary Committee's sixteen-month investigation into digital markets recommended mandating data portability through legislation requiring data portability under competition law. Although both the U.S. Supreme Court and the European Court of Justice have yet to adopt the doctrine, their discussions of related principles and lower court decisions imply some clear prerequisites for it to apply. Both jurisdictions require that the resource to which competitors are demanding access be *essential* or *indispensable*, in that the competitor is practically unable to duplicate it or obtain it from another source.² U.S. law also requires the absence of a regulatory regime through which the competitor could obtain access to the resource and the feasibility of providing access.³ European law limits mandating access to "exceptional circumstances," further requiring that refusal to provide access excludes effective competition and prevents the appearance of a new product or new technological development, which precludes the use of data portability to provide me-too products.⁴

These requirements represent important prerequisites that must be satisfied before mandating data portability under competition law. In addition, the fact that options for data portability remain virtually unused despite being mandated by GDPR and CCPA and offered by leading digital platforms, such as Google and Facebook, suggests that the implementation of data portability may present important practical considerations that must be taken into account.

² Case T-201/04, *Microsoft v. Commission*, 2007 E.C.R. II-3601, 3725 ¶ 328; 3B PHILLIP E. AREEDA & HERBERT HOVENKAMP, *ANTITRUST LAW* ¶ 773b2, at 242-43 (3d ed. 2008).

³ *Verizon Commc'ns Inc. v. Law Offices of Curtis V. Trinko*, 540 U.S. 398, 411 (2004); *MCI Commc'ns Corp. v. AT&T Co.*, 708 F.2d 1081, 1133 (7th Cir. 1983).

⁴ *Microsoft*, 2007 E.C.R. at II-3726 ¶¶ 331-332, II-3818 ¶ 647.

II. ALL DATA ARE NOT CREATED EQUAL: STRUCTURED VS. UNSTRUCTURED DATA

An essential prerequisite for mandating data portability under competition law is the inability for the competitor to duplicate the data or obtain them from another source. To date, discussions have largely treated data as a monolithic phenomenon without drawing any distinctions among particular types of data and their different uses. A proper evaluation of the opportunities for self-provisioning and the availability of similar data from third parties requires a better understanding the precise types of data to which access is sought and their uses.

For example, although advocacy rhetoric tends to talk about “big” data, the trade press repeatedly emphasizes that size is not the only thing that matters. Gartner famously articulated in 2001 that data consists of 3 Vs: Although volume is an important characteristic of data, so are its velocity (rate of change) and variety (differences in type and source). From there, other commentators have expanded the list, with one consulting firm listing as many as 42 Vs.⁵ Although this framework has yet to be subject to rigorous academic analysis, the basic point about the multidimensionality of data is clear enough.

There is another more important distinction that has yet to arise in competition law circles but has begun to attract attention among academics: the distinction between structured and unstructured data. *Structured data* are collected intentionally to inform a specific model. Examples include traditional column-row databases that record names, dates, addresses, and transaction histories, which represent the type of data with which people are most familiar. *Unstructured data* are collected incidentally and used to inform emergent models. Examples include video, audio, social media feeds, photos, and sensor data. *Semi-structured data*, which are structured data used to inform other, emergent models, represent an intermediate case. A common example of semi-structured data is email analyzed for purposes other than person-to-person communications.

Structured and unstructured data serve very different purposes. The user-facing side of online transactions (such as search results or purchase recommendations) tends to be based on structured data. The profiles used to support advertising incorporate significant amounts of unstructured data.

More importantly for purposes of competition law, structured and unstructured data have different economic characteristics relevant to the essential facilities doctrine. Empirical studies have shown the scale economies with respect to structured data to be modest enough that relatively small competitors should be able to achieve them on their own without gaining access to the resources of others. That fact would tend to vitiate claims that the facility is essential. The scale economies for unstructured data are more significant, but the number of alternative sources of the demographic and other information used to create advertising profiles are legion. Indeed, industry observers suggest that more than 80 percent of all available data is unstructured data.

These considerations suggest some key factual propositions regarding scale economies and substitutability that must be established by any antitrust court faced with a request to mandate data portability. They also underscore the importance of considering such requests in the context of specific types of data and not lumping all forms of data into the same bucket.

⁵ Tom Safer, *The 42 Vs of Big Data and Data Science*, ELDER RESEARCH (Apr. 1, 2017), <https://www.elderresearch.com/blog/42-v-of-big-data>.
CPI Antitrust Chronicle November 2020

III. THE SIGNIFICANCE OF REGULATORY ALTERNATIVES

As noted earlier, different legal regimes draw different inferences from the existence of a regulatory regime that could provide access to the resource in question. U.S. law holds that the presence of a regulatory regime through which competitors can obtain access renders the essential facilities doctrine inapposite, while the recent decision of the German Federal Court Justice in the Facebook case ruled the presence of related relief under privacy law does not foreclose actions under competition law.⁶ This distinction accord with the U.S. tradition of treating competition law and regulation as substitutes and the European approach of treating them as complements.⁷

These jurisprudential differences suggest that each regime may accord different effect to the fact that CCPA and GDPR already provide for the right of data portability, with the U.S. regarding the existence of a regulatory mechanism for compelling sharing rendering access under the antitrust laws nonessential. The fact that the relevant European privacy law did not cover the precise conduct in question in the German Facebook case does leave some room for a different outcome with respect to data portability.

IV. THE TENSION BETWEEN PORTABILITY AND PRIVACY

One of the central raised concerns at the FTC hearing was the possibility that data portability may impair efforts to protect consumer privacy. As an initial matter, any data portability regime must take great care to verify that the person requesting the data is actually the data subject. Any failure to verify the person's identity risks giving unauthorized actors access to their personal information.

In addition, one of the fundamental cornerstones of U.S. privacy law is notice and consent, in which the parties enter into an agreement as to the permitted uses of any data collected. This makes the enforcement of promises made in privacy policies a matter of contract. While data subjects have sufficient privity of contract with the firm that collects their data in the first instance to have standing to sue them for any breaches of that agreement, that contract does not provide any enforceable rights against third parties that obtain the information through data portability. Thus, exercises of data portability carry some risk of weakening privacy protections.

The same problem arises with respect to re-identification of data. The most common way to de-anonymize a dataset is by correlating it against an identified dataset. The best practices identified by the FTC's 2012 report require holders of de-identified data to "(1) take[] reasonable measures to ensure that the data is de-identified; (2) publicly commit[] not to try to re-identify the data; and (3) contractually prohibit[] downstream recipients from trying to re-identify the data."⁸ The reliance on contract to protect against de-identification again places data portability in tension with privacy by potentially limiting the data subjects' rights against firms that use data portability to obtain access to their personal information.

V. IMPLEMENTATION CHALLENGES

Beyond the conceptual challenges discussed above, mandating data portability under antitrust law would raise significant practical challenges as well. It would require establishing new systems for ordering and provisioning requests for data. It would also raise significant presume to mandate data compatibility as well.

Ordering and provisioning: Lower U.S. decisions also predicate the essential facilities doctrine on the feasibility of providing access.⁹ As the Supreme Court noted, services that are not already offered as separate products "exist only deep within the bowels" of the company, which means that access mandates would force companies to design and implement new systems to make that access possible.¹⁰

6 Compare *Trinko*, 540 U.S. at 411-13, with *BGH* June 23, 2020, KVR 69/19 ¶ 126, <https://juris.bundesgerichtshof.de/cgi-bin/rechtsprechung/document.py?Gericht=bgh&Art=pm&Datum=2020-6&nr=109506&linked=bes&Blank=1&file=dokument.pdf>.

7 Compare *Pac. Bell Tel. Co. v. linkLine Commc'ns, Inc.*, 555 U.S. 438 (2009), with Case C-280/08P, *Deutsche Telekom v. Commission*, 2010 E.C.R. I-9555.

8 FED. TRADE COMM'N, PROTECTING CONSUMER PRIVACY IN AN ERA OF RAPID CHANGE: RECOMMENDATIONS FOR BUSINESSES AND POLICYMAKERS iv, 21 (2012), available at <https://www.ftc.gov/sites/default/files/documents/reports/federal-trade-commission-report-protecting-consumer-privacy-era-rapid-change-recommendations/120326privacyreport.pdf>.

9 *MCI*, 708 F.2d at 1133.

10 *Trinko*, 540 U.S. at 410. *But see* Case C-418/01, *IMS Health GmbH & Co. OHG v. NDC Health GmbH & Co. KG*, 2004 E.C.R. I-5069, I-5084 ¶ 43 (noting that service need not be marketed separately to fall within essential facilities doctrine).

In short, a data portability mandate would require firms to establish new systems for ordering and provisioning request for data. Court considering implementing such mandates will need to decide how much metadata will be disclosed. If the data are substantial, the process may take some time and require considerable resources to transfer the data. Past experience has shown that antitrust courts will have to supervise an increasing scope of the business relationship when the quality of the product varies and when the interfaces are complex.¹¹

Compatibility: The fact that firms base their operations on different and often incompatible data structures means that simply mandating data portability will require competitors to fit the proverbial square peg into a round hole. To promote competition, the data that is being shared must be compatible.

The need for compatibility suggests two natural solutions, both of which have considerable drawbacks. The first is reconfiguration of the data, most likely by the competitor who requested them. The problem is that enterprise-grade databases are often prohibitively expensive to reconfigure. In addition, any process of conversion or translation runs the risk of introducing errors.

The second is standardization. Any form of standardization of data formats necessarily structures interactions in ways that can limit the functionality of these systems. Requiring that a firm use a particular data structure thus inevitably has a direct impact on innovation.¹² In addition, antitrust officials have become increasingly concerned that standard setting processes may be vulnerable to strategic behavior.

VI. CONCLUSION

Some of the participants in the FTC workshop have somewhat ingenuously pointed to data portability as low hanging fruit in terms of antitrust remedies for big tech firms compared with the challenges posed by more complex remedies such as interoperability. The need to differentiate between structured and unstructured data; the relevance of alternative regulatory mechanisms for obtaining access to data; the potential for portability to weaken privacy; and implementation challenges relating to ordering, provisioning, and compatibility reveal that data portability is not the panacea for which some would hope. Instead, a closer examination confirms that such cases require the type of nuanced, fact-specific inquiry that characterizes classic antitrust analysis and the implicit limitations of the still provisional essential facilities doctrine.

¹¹ *Trinko*, 540 U.S. at 414; 3B AREEDA & HOVENKAMP, *supra* note 2, ¶ 774e, at 275-79.

¹² Christopher S. Yoo, *When Antitrust Met Facebook*, 19 GEO. MASON L. REV. 1147, 1155 (2012).



CPI Subscriptions

CPI reaches more than 35,000 readers in over 150 countries every day. Our online library houses over 23,000 papers, articles and interviews.

Visit competitionpolicyinternational.com today to see our available plans and join CPI's global community of antitrust experts.

