



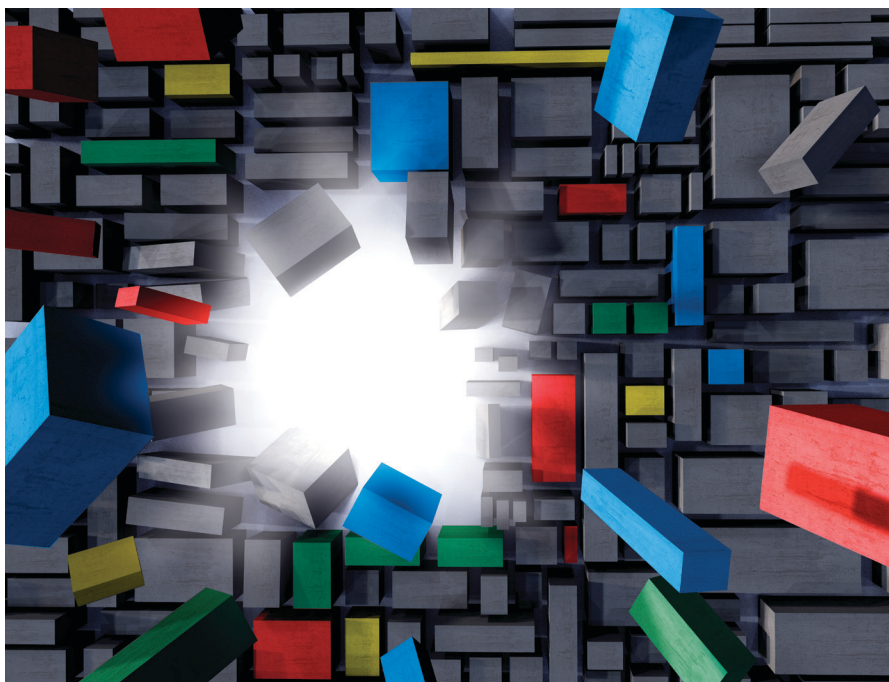
Viewpoint

Breaking Up a Digital Monopoly

How to decompose a vertically integrated digital monopoly to enable competitive services based on a shared data structure.

THE DOMINATING POWER of today’s global data monopolies—most prominently Google, Facebook, and Amazon—has alarmed people around the world. Governments are seeking ways to rein in such monopolies and establish reasonable conditions for competition in the services they offer. Their business models (such as targeted advertising), also raise major issues of personal security and privacy.⁹ Thus, measures that control their tendencies toward monopoly may help to address the threats they pose to civil and political liberties. In this Viewpoint, we propose a regulatory strategy that addresses the *naturally monopolistic* nature of these services by isolating the *core acquired data* collection and management functions. Acquired data is data derived from the discourse of society at large so the public retains a legitimate ownership interest in it. As described in this Viewpoint, our proposal requires companies to compete by innovation rather than through monopolistic control over data.

The dominant data platforms can arguably be characterized as natural monopolies within their respective type of service. According to Richard Posner’s classic account, “If the entire demand within a relevant market can be satisfied at lowest cost by one firm rather than by two or more, the market is a natural monopoly, whatever the actual number of firms in it.” He goes on to say that in such a market the



firms will tend to “... shake down to one through mergers or failures.”⁶ Among producers, the costs of entry, such as necessary infrastructure investment, leads to large economies of scale when there are few producers, and this tends to give an advantage to the largest supplier in an industry. This phenomenon is captured by the concept of subadditivity, which is the basis for the modern theory of natural monopoly (see the accompanying sidebar). Among consumers, services that benefit from strong network effects also tend to dominate over time.⁷ Familiar examples of natural monopolies include

public utilities such as water services, the electricity grid, and telecommunications.

In the case of the *acquired data monopolies* that societies struggle with today, there is a second structural factor. Such services are based on the collection and management of information intended for a wide or unrestricted audience. They reap huge profits by exploiting a remarkable ability to monetize this data at scale. We call the digital content that these platforms build on “acquired data” to indicate it is either collected from unrestricted public sources (for example, Web pages or

street cameras), or that it is provided by users who relinquish ownership in order to have it managed and distributed to others. We introduce the term acquired data to help distinguish it from surveillance data, which is collected from users without their explicit consent or agreement. An example of surveillance data is user information derived from keystrokes during data entry or from tracking of online behavior using third-party cookies. A third category is inferred data, which can be derived from published content through statistical correlation. An example of inferred data is determining the author of an anonymous article through their use of words. Importantly, and in contrast to information collected solely through surveillance or inference, sharing what we term acquired data does not necessarily raise privacy or security issues since it is by definition either intentionally made public or submitted for publication.

We contend the public retains a legitimate ownership interest in acquired data in spite of the user having possibly assigned their rights to the distributor. This is akin to the idea that a contract entered into under duress is not necessarily enforceable. When the means of distribution is monopolistically controlled, users having few other means to express themselves publicly may be coerced into accepting unfair conditions. Moreover, while there is no explicit cost to users who hand over their content, there are implicit ones. One implicit cost is required juxtaposition of content with advertisements sold by the distribution service. Another type of implicit cost may be providing access only within a pay-to-play “walled garden.”

Our proposal rests on the notion that distribution of public discourse and other acquired content serves the common good of the community of content providers and consumers. Treating it as a private asset of the distribution service does not.⁵ The main goal of decomposing acquired data monopolies is to ensure the distribution of such content at reasonable cost, both implicit and explicit. This may require overcoming the naturally monopolistic nature of such a service through regulation.

Our proposal rests on the notion that distribution of public discourse and other acquired content serves the common good of the community of content providers and consumers.

How Can a Digital Platform be Decomposed?

Broadly speaking, service corporations that assert monopoly power do so along two different axes: either they integrate “horizontally” by creating a portfolio of related services and buying up competitors; or they integrate “vertically” by increasing their direct ownership of various stages of production and buying up suppliers. As with Standard Oil at the beginning of the 20th century and AT&T at its end, one obvious way to decompose today’s online conglomerates is “horizontally,” for example, by separating Meta (Facebook, Instagram, WhatsApp) or Google (including Web search, Web analytics, cloud storage and Gmail) into smaller and more specialized companies.

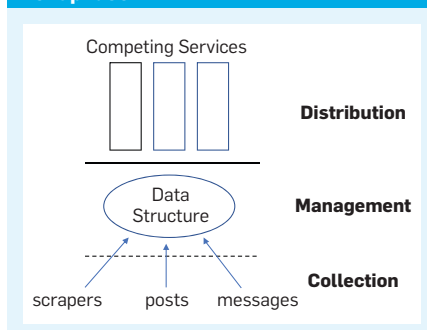
Horizontal decomposition has a mixed track record. In the early 1980s the Bell Telephone System was broken up into the AT&T Corporation, providing long-distance lines and equipment, and the seven local Bell operating companies. However, subsequent reconsolidation within the telephone industry has returned to a small number of companies. This example shows how legislative remedies that do not address underlying technical and economic realities may ultimately fail to create a broadly competitive market.

By contrast, “vertical decomposition” splits a production into components that have different structure, makeup, and economic characteris-

tics. Looking at acquired data services, we note they generally have two distinct phases (see the accompanying figure): In the collection/management phase, material is acquired from content producers (that is, end users of the service) or collected from public sources.⁸ A data structure that organizes it is constructed, managed, and accessed through backend APIs. In the distribution phase, companies create different services for their users and customers that access this data structure and its content (for example, user-initiated search, per-member “news” feeds, and so forth). It is in the distribution phase that the most aggressive monetization strategies are pursued. In practice, the two phases often operate concurrently through continuous update of the acquired data structure.

We argue that one important aspect of the tendency to natural monopoly among acquired data services is the generic nature of the acquired data structure created in the collection and management phase.³ Multiple Web crawlers indexing the same publicly visible Web pages, or cameras surveying the same streets, will generate broadly equivalent data structures. Another factor is the cost of entry into the market for this phase. In the case of Web search, for example, the content acquisition phase accumulates in a data structure called a “search index” built by Web crawlers. Google reports their search index has on the order of 10^{18} entries, which is certainly many petabytes in storage size, making the cost of creation, maintenance and access a barrier to entry. Similarly, Facebook builds a massive “social graph”

Vertical decomposition. An acquired data monopoly is structured as two phases: collection/management and distribution. Tendencies toward natural monopoly are concentrated in the collection/management phase.



from contributed posts and other interactions with members who voluntarily exchange their “content” for unspecified distribution services. In the case of a social media service like Facebook, another salient factor lies in the network effects of having a single large social media provider.⁸

Due to the two-phase structure of acquired data services, the tendency toward natural monopoly can be isolated in the initial collection and management phase by vertically decomposing the service. This would allow multiple distribution services based on the same data structure to compete on their merits. Competition would then stimulate differentiation between distribution phase services that rely on acquired content. Service providers building on top of a shared data structure in this way could not leverage a monopoly over acquired data, and would be forced to generate value competitively. Current efforts in this space include Mastodon open source social media and DuckDuckGo in privacy-preserving search.

At least one effort to promulgate such a strategy has already been initiated. In 2020, software developer Zack Maril founded the Knuckleheads’ Club with the goal of opening up Google’s search index as a public utility.⁴ Although different acquired data monopolies may acquire their data in different ways and build on different types of data structures, they are vertically integrated in the same way. So our

Some have pointed to interoperability as the key to improved competition in acquired data services.

proposal is to apply the same fundamental strategy to each of them.

Scalable Sharing of a Weak Common Service

Some have pointed to interoperability as the key to improved competition in acquired data services.² Adopting a shared data structure for acquired data would enable interoperability in the definition of higher layer services based on it. But since a standard that aims to deliver enduring interoperability must gain wide acceptance and remain useful as technologies and environments evolve, agreeing on such a standard is a central challenge for creating infrastructure at scale. Based on historical successes such as the Internet protocol suite and the POSIX kernel API, it is widely believed the key to defining a common service interface that can support many specific value-added high layer services is the famil-

iar “hourglass model,” with the common interface as its “narrow waist.” The fundamental question is: How is an interface to be designed that is minimal yet still sufficient to the purpose?

As we have shown elsewhere, the *Hourglass Theorem* is a design principle that can guide the development of a successful service stack that has this structure.¹ The Hourglass Theorem says that if you make the service interface you are designing *logically weaker*, you will increase the class of possible lower-level services that can be used to support it. However, you will also decrease the class of possible applications it can support. In the case of the POSIX operating system interface, weakness means that certain classes of applications cannot be supported without extensions to the kernel interface. Examples are those which require real-time scheduling or parallel file access. In the case of the Internet, weakness means that applications requiring quality of service, highly reliable, or undetectable data transmission cannot be supported without modifications to the architecture.

But in both cases, application communities have been willing to accept the limitations imposed by the standard in order to obtain a service that can be implemented across an incredible diversity of environments, and therefore provide interoperability. These environments range from the smallest personal and IoT devices to Cloud virtual machine clusters and exascale supercomputers. The Hourglass Theorem tells us this trade-off is inevitable when defining a common interoperable service, and thus offers insight into the likelihood of its *voluntary* widespread adoption.¹

When it comes to the vertical decomposition of a given acquired data monopoly, we seek a data structure that can serve as a standard for interoperable data exchange. The Hourglass Theorem tells us that to achieve this goal, this interface must support a set of target applications (the upper bell of the hourglass) deemed necessary for success. However, given this constraint, the interface should be as weak as possible.

This weakness (the thinness of the waist) maximizes the class of possible environments that can support it (the

The Theory of Natural Monopoly*

The modern view of natural monopoly (also called the “New Learning”) is based on the concept of subadditivity. That is, the salient feature of natural monopoly is a cost function that is subadditive.

William Sharkey, in his book *The Theory of Natural Monopoly* defines subadditivity as follows:

If q_1, q_2, \dots, q_k are output bundles that sum to q , then a single firm is superior on efficiency grounds to a multifirm industry if the following condition holds:

$$C(q) < C(q_1) + C(q_2) + \dots + C(q_k) \quad (1)$$

$C(q_i)$ can be interpreted as the cost of producing commodity bundle q_i . If inequality (1) holds, then a single firm can *jointly* produce bundles q_1, q_2, \dots, q_k more cheaply than if the bundles were produced separately, or if they were produced by two or more firms.

Think of subadditivity as an extension of the concept of economies of scale to the multiproduct case.

*From Chris Brown, “The Theory of Natural Monopoly,” *Antitrust, Economic Regulation, and Social Regulation*, (Mar. 29, 2008); <https://bit.ly/3HgDCOD>

lower bell). For example, metadata collected by a Web crawler consists of a number of attributes that go into the generation of a response to a search query. Such attributes include word and link counts. A substantial fraction of the cost of constructing and maintaining a Web search index goes into building it from this acquired data. A certain level of query processing and content matching based on this index can be done through a weak API. However, the natural language processing that precedes such a user query and the ordering of results according to a ranking algorithm could be proprietary to a particular search provider.

A similar analysis could apply to a social media platform, with the data structure being the social graph that holds posts and discussion attributes, such as views and likes. The creation of groups and the generation of feeds would then be implemented as a higher-level supported service.

Thus, in the context of shared services for acquired data, a stronger service interface restricts the possible implementations in a number of ways that can reduce competition and increase the tendency toward monopoly. In the context of search, any requirement beyond the best effort collection and dissemination of information can make the operation of the service more expensive and thus create barriers to entry. Well-intentioned requirements, such as “the right to be forgotten” or the need to block obscene or hateful content, are examples. Enhancements for usability, such as having to deliver results in page rank order, place a large burden on the operator to collate and tally such metrics. Another class of restrictions includes those that can only be supported by certain proprietary platforms, such as a specific instruction set architecture. If the proposed common service requires that requests are replied to within a certain time-frame, or that expensive or difficult image analysis be performed, that can likewise create barriers to entry.

Another way that logical strength can be used to create barriers is for the common interface to require information or services that are proprietary to the service operator. If, for example, the results of a search or feed creation algorithm are a sequence (rather than

Historically, companies faced with a standardized alternative to their core service have responded defensively.

an unordered set) then the ordering can be a proprietary mechanism for adding value. A minimally strong service would separate the “raw” service that returns a set from the “cooked” service that returns a list. However, to ensure the common service enables the implementation of ordered lists, it may be necessary to also expose some metadata.

Conclusion

Historically, companies faced with a standardized alternative to their core service have responded defensively. Adopting a more generic data management layer that can support competitive forms of distribution turns their proprietary asset into a commodity. Classical examples are the telephone system, and local area network vendors. In these cases, POSIX and the Internet were initially dismissed and adoption refused. Some famous cases resulted in the near or complete extinction of the companies that resisted them. In other cases, these standards were adopted or similar proprietary versions created to maintain market segmentation. The largest corporations in the world now base their information ecosystems on such standards.

Our proposal is to break up acquired data monopolies vertically, leveraging the hourglass design principle to create durable but weak common services that can be regulated as public utilities. Whether our society can come to terms with social media and acquired data depends on our willingness to assert a public interest in public discourse, as well as the slow but hopefully inexo-

nable advantage of scalable standards over monopolistic strategies.

Since we are not lawyers or professional policymakers, we shall not attempt to offer an authoritative legal or political theory under which the companies that now profit from monopolistic control over acquired data could be convinced or compelled to share it. Traditionally, monopolies are criticized for imposing high costs on their customers, but in this case the costs are implicit in the capture of attention and the profits extracted indirectly through means such as advertising.

The fact that control of the digital environment invades so many aspects of modern life makes these indirect costs more burdensome, if less obvious. The information technology industry has successfully used organizations such as the Internet Engineering Task Force and the World Wide Web Consortium to achieve consensus around open interoperable standards when it serves their needs. The key element that may require regulation is to dispel the notion that data acquired through mechanisms that lead to monopoly are the private property of the agent that collects it. ■

References

1. Beck, M. On the hourglass model. *Commun. ACM* 62, 7 (July 2017); DOI:10.1145/3274770.
2. Doctorow, C. Competitive compatibility. *Commun. ACM* 64, 10 (Oct. 2021), 26–29; DOI:10.1145/3446789.
3. Ducci, F. *Natural Monopolies in Digital Platform Markets*. Cambridge University Press (2020).
4. Maril, Z. Knuckleheads' Club; <https://bit.ly/43Nj9tZ>
5. Mazzucato, M. Let's make private data into a public good. *MIT Technology Review* 121, 4 (2018), 74–75.
6. Posner, R.A. Natural Monopoly and Its Regulation 21 *Stanford Law Review* 592 (Feb. 1969). *Journal of Reprints for Antitrust Law and Economics* 5, 335 (1973).
7. The NFX Team 2022. The Network Effects Bible; <https://bit.ly/3KUaT2X>
8. U.S. House Judiciary Committee 2020. Investigation of Competition in Digital Markets: Majority Staff Report and Recommendations. Subcommittee on Antitrust, Commercial, and Administrative Law Committee on the Judiciary, U.S. House of Representatives, Washington, D.C.; https://bit.ly/USHJC_Competitive_Digital_Markets
9. Vardi, M.Y. ACM, ethics, and corporate behavior. *Commun. ACM* 65, 3 (Mar. 2022); <http://dx.doi.org/10.1145/3516423>

Micah D. Beck (mbeck@utk.edu) is an associate professor at the Department of Electrical Engineering and Computer Science, University of Tennessee, Knoxville, TN, USA.

Terry R. Moore (tmoore@icl.utk.edu) is an associate director (retired) at Innovative Computing Laboratory, University of Tennessee, Knoxville, TN, USA.

The authors gratefully acknowledge the contribution of Elaine Wenderholm.

Copyright held by authors.