Power, Privacy and Personalization in Digital Commerce

Nirvikar Singh*

Rajiv Sunkara**

Simon Yencken***

March 26, 2021

Abstract

This paper reviews emerging issues with respect to market power and user privacy in digital commerce, and implications for the use of personalization by providers of products and services. It argues that advances in various forms of AI, machine learning, and related software tools allow for the effective use of digitally-generated behavioral data, both live and historical, as well as addressing concerns about market power and digital privacy. The contribution of this paper is to integrate the treatment of personalization based on behavioral data with concerns about privacy and market power in digital networks, to illustrate effective solutions to these problems, and bring out the managerial implications of this approach.

JEL Codes: L12, L50, L86, O30

Keywords: behavioral data, personalization, artificial intelligence, market power, data privacy, digital networks

---

* Professor of Economics, Department of Economics, University of California, Santa Cruz, 1156 High Street, Santa Cruz, CA 95064, USA, boxjenk@ucsc.edu (corresponding author)

** CTO, Fanplayr, 725 Alma Street, Palo Alto, 94301, USA

*** CEO, Fanplayr, 725 Alma Street, Palo Alto, 94301, USA

Power, Privacy and Personalization in Digital Commerce

Introduction

If we begin with the creation of the ARPANET in 1969, we have just crossed the half-century mark of what might be termed the “digital network age” (Singh, 2020). This period has been marked by transformative developments in society, politics and economics, influenced by digital communication networks. In addition to general information gathering and consumption of entertainment and news content, the Internet and Worldwide Web have become significant local, national and global sites for social and professional interactions, commercial activities, and political speech. The COVID-19 pandemic has only underlined the importance of digital networks for how we work, learn and play.

Aside from the benefits of enormous reductions in the cost of digitally-enabled communications and transactions, several issues have arisen, including challenges of unchecked dissemination of misinformation and problematic content. Another problem that is less obviously alarming, but has been attracting steadily increasing concern, is the erosion of privacy for individuals who use digital networks for a wider and wider range of activities. Both the spread of problematic information and content, and the erosion of privacy, draw attention to the role played by giant technology firms that provide platforms and tools for information flows and storage. There is increasing analysis of the novel forms of market power that are being exercised in the context of digital networks.

Market power can be expressed in different forms, depending on the nature of the technology firm and its services. Online retail services (e.g., Amazon), information and content (e.g., Alphabet) and social networks (e.g., Facebook) each have somewhat different sources of scale advantage. But in each case, the control of data and impacts on privacy are an important concern associated with this market power, in addition to the traditional, more straightforward issue of elevated prices for products and services. The use of individual data can support reductions in competition and higher prices, though it can also allow for better serving those individuals. This is the potential benefit of “personalization.” Often, the ethical issues revolve around informed consent, or putting the collection and use of some kinds of data off limits, making this not just a question of unfair transfers of money from buyers to sellers.[1] New kinds of digitally-generated data add to the complexity of the issue, since relatively anonymous, transitory data[2] has different privacy implications than demographic and other personal-identification data.

This paper discusses the interrelated issues of market power and data privacy in digital networks, and how new, more privacy-respecting, approaches to the use of individual data can preserve the benefits to buyers and consumers of services that come from better understanding of their wants and needs. This approach can also have implications for the exercise of market power in digital networks. New software tools that come under categorizations such as Artificial Intelligence (AI), Advanced Analytics (AA), and machine learning are central to this approach. The rest of the paper is structured as follows. The next section reviews analyses of market power and data privacy in the context of digital networks. The third section discusses the use of new software tools to manage digital behavioral data to better serve buyers and consumers, while being more privacy-respecting. This also includes a modified concept of personalization, which need not be tied to fixed or ethically sensitive personal characteristics. The fourth section provides a concrete illustration of how a focus on behavioral data can be implemented, as well as of potential managerial and bottom-line impacts. The fifth and final section concludes with a summary and a discussion of possible developments over the next few years.

Market Power and Data Privacy

Global trends in industry structures indicate increasing differences in size and productivity that are much broader than just the case of the digital giants that draw everyday attention.[3] These changes are associated with higher price-cost markups for larger firms, which can be an indicator of market power. Weaker anti-trust enforcement, for which there is some evidence, is consistent with the hypothesis of increased market power as a contributing factor. However, globalization and technological innovation are possibly more important explanators of the emergence of "superstar" firms that outstrip the rest of their industries. Even if these are not technology firms,

[1] Of course, the dominance of large technology firms can also lead to the demise of small businesses, with a knock-on effect on their individual owners, in addition to direct impacts on buyers.
[2] This characterization of data is closely related to the somewhat broader term "behavioral data." These issues of nomenclature are discussed later in the paper.
[3] The following brief discussion is based on the detailed survey by Van Reenen (2018), who provides numerous source references.

they can take advantage of digital technology and digital networks. So, a Walmart can benefit from these innovations as well as an Amazon.[4]

Beyond more traditional sources of market power, digital networks provide size advantages, which in turn can lead to market power. This is the concern about digital giants like Amazon, Alphabet and Facebook, among quite a few others.[5] What is the link from networks to size advantage? The basic benefit of a large network arises from an individual member of the network being able to have more interactions the larger the network. This property was understood prior to the Internet, since it is a feature of all telecommunication networks (Rohlfs 1974; Katz and Shapiro 1985). The network effects here are termed "direct," since any individual potentially gains more from joining a larger network. Digital networks such as Facebook and LinkedIn also have this kind of property.

On the other hand, firms such as Uber or eBay operate in two-sided networks, made up of drivers and riders, or sellers and buyers. Here, the network effects are indirect: "the number of users on one side of the market attracts more users on the other market side."[6] These markets are known as "platform" markets, since an eBay or an Uber provides the software platform for users to interact.[7] There is an analogy, perhaps, to the development of department store chains, which created replicable store layouts as "platforms" for bringing consumers and producers together. The difference, of course, is that scaling up that model requires land and buildings in many locations, which is much costlier than adding web servers and other software to handle more online traffic.[8]

[4] Walmart was, of course, a pioneer in creating digital supply chains, using proprietary technologies, well before the emergence of the Internet.

[5] To a considerable degree, Apple is an outlier among the digital giants, relying much less on advertising over digital networks, or network economies in general, instead creating brand loyalty based on superior design of hardware and software, and an appealing image.

[6] From Haucap and Heimeshoff (2014), p. 51: notable earlier analyses are Rochet and Tirole (2003) and Evans and Schmalensee (2007). A recent comprehensive analysis of platform market power is Bamberger and Lobel (2017).

[7] Non-market networks such as Facebook are also platforms, of course. The case of LinkedIn can be viewed as a hybrid, since it provides a platform for general professional networks, but also has some features of a two-sided platform market of employers and employees. The Google search business of Alphabet, by far its most profitable component, is also a two-sided market, bringing together information-seekers and advertisers. In this case, the audience for the advertisers is narrower than the universe of information-seekers, but that overall set is so large that the business is very successful.

[8] Of course, this is why Amazon has been able to grow more rapidly than Walmart, which took many decades to reach its dominant position in pre-Internet retailing in the US.

To summarize, digital networks can support new sources of economies of scale, emerging from the demand side as well as the more traditional supply or production side. These "network economies" can lead to market power that is associated with size, along with efficiencies in supply that come from conventional scale economies. An even more novel source of market power, however, comes from the gathering and analysis of user data, especially in the case of households and individuals.

The analysis of consumer data has a long history, of course. The ability to digitally store and analyze demographic and purchasing information led to a quantitative revolution in marketing by the 1980s, if not earlier. Segmentation and targeting of individuals and households became more sophisticated, using statistical and other quantitative tools. One example of pre-Internet developments was the introduction of loyalty programs by grocery stores, with discounts for members incentivizing the revelation of who was buying what, and when. Frequent flyer programs were another early example, though in their case, lock-in through cumulative earned rewards may initially have been more important than segmentation and targeting. Both these examples involved first-party data collection from customers, with some degree of consent and transparency. The expanded use of credit cards led to "third-party" data collection on a massive scale, in the sense that a credit card user is a customer of the card provider, but the detailed spending data was more valuable to the sellers of goods and services.

As we know, digital networks have led to an explosion of such third-party data collection, often without transparency or well-informed consent. Search behavior on Google with clearly marked sponsored ads was initially an innocuous process, but intermingling of sponsored and non-sponsored search results, and manipulation of the ordering of results eventually emerged as concerns, alongside the collection of related data from web-based email platforms, and other services. In the case of social networks like Facebook, the collection of personal data to use for targeted advertising became even less transparent, since it occurs in the context of a wide range of social interactions and information sharing.[9,10]

These developments are well known. What do they imply for market power? Nuccio and Guerzoni (2019) put data gathering at the center of market power for digital platforms (p. 313): "The competitive advantage of these companies relies on the data they can gather from their users and customers in order to fuel predictive models capable of probabilistically determining preferences and purchase behaviours." Of course, as discussed earlier in other contexts, the threat of market power – in this case from access to user data – is intertwined with the ability to serve users better by understanding their wants better. More explicitly, user data can create more value as well as enhancing the capture of value. Many other analyses of the market power associated with the collection and analysis of data by digital platforms acknowledge other sources of market power, as well as the different levels of potential for exploiting data, depending on the nature of the platform (e.g., Facebook vs. Twitter or Spotify).[11]

A major question in all of the analyses of market power created by control of digital data is the degree of excessive extraction of value from users by the large digital platforms. However, there are also more subtle issues of privacy rights and more general concerns about the ownership and ethical use of personal data. It seems fair to say that a combination of these broader ethical concerns with more traditional market power issues associated with the use of data by large

[9] As noted earlier, Apple is an exception among the digital giants, and its CEO recently weighed in with a powerful criticism of how data is being used by some of his company's business rivals (Bariso, 2020): "Technology does not need vast troves of personal data stitched together across dozens of websites and apps in order to succeed. Advertising existed and thrived for decades without it, and we're here today because the path of least resistance is rarely the path of wisdom. If a business is built on misleading users on data exploitation, on choices that are no choices at all, then it does not deserve our praise. It deserves reform." Cook also explicitly connects a business model based on indiscriminate data collection to problems of misinformation plaguing contemporary society: "We should not look away from the bigger picture and a moment of rampant disinformation and conspiracy theory is juiced by algorithms. We can no longer turn a blind eye to a theory of technology that says all engagement is good engagement, the longer the better, and all with the goal of collecting as much data as possible."

[10] We do not go into it in this paper, but the use of home assistants such as Alexa, and the Internet of Things, with networked thermostats, home security devices, refrigerators, wearable devices, and automobiles, also represent additional sources of privacy erosion and market power – the latter especially in the case of technology giants such as Alphabet and Amazon. Two examples in a burgeoning literature on this topic are Weber (2015) and Maple (2017).

[11] In addition to Nuccio and Guerzoni (2019), Haucap and Heimeshoff (2014), and Bamberger and Lobel (2017), Graef (2015) and Prado (2020) are useful references. Each of these analyses provide additional references. Of course, this literature is growing rapidly.

digital platforms is what is driving various forms of privacy regulation of data generated by the use of digital networks.

We do not analyze the complexities of these issues,[12] merely note that they are already having an impact on digital commerce. Three examples of the new emphasis on protecting privacy of personal data and informed choice by users of digital networks are the European Union, Japan, and the state of California in the US. De Groot (2019) summarizes the new European Union regulations, including provisions requiring explicit consent and anonymization, among other individual protections, along with stiffer penalties for compliance failures, and global implications. The EU's own site (eugdpr.org) is very blunt about the changes: "The EU General Data Protection Regulation (GDPR) is the most important change in data privacy regulation in 20 years. The regulation will fundamentally reshape the way in which data is handled across every sector, from healthcare to banking and beyond." Japan amended its digital privacy laws in a manner that moves them close to the EU GDPR.[13] California's Consumer Privacy Act, which took effect on January 1, 2020, has been described as "the toughest data privacy law in the U.S."[14]

The development of digital commerce has arguably been unbalanced, driven by traditional economies of scale, network economies, and competitive advantages based on large-scale collection of digital data. The results have been the rapid rise of new digital platform firms, and following on that, intertwined concerns about market power and privacy of personal data have led to significant new regulations. This is the situation in which we proceed to describe emerging approaches to handling digital data, which are adaptable to the increased requirements of digital data privacy, and which can provide firms other than the platform giants with possibilities to thrive.[15]

---

[12] For some examples that discuss the legal, ethical and technical issues involved in data privacy laws and regulations, see Sarathy and Robinson (2003), Rice and Sussan (2016), Sisk (2016) and Gaus (2017).

[13] A translated version of Japanese digital privacy laws is available at: http://www.japaneselawtranslation.go.jp/law/detail/?id=2781&vm=04&re=01.

[14] See: https://www.npr.org/2019/12/30/791190150/california-rings-in-the-new-year-with-a-new-data-privacy-law.

[15] We may note that there are opposing trends as well. For example, Chinese e-commerce firms offer a range of services bundled with shopping, including entertainment, social networking, messaging and payments. This is a model which enhances the concentration of digital data, and potentially increases market power. Some analysts argue that this model will transfer well to Western markets: see *The Economist* (2021). Much will depend on consumer and regulatory responses, and it is possible that different approaches will co-exist.

Working with Behavioral Data

In the context of digital networks, behavioral data covers a range of possibilities: browsing characteristics such as total time on site, speed, page views, and exit intent, in addition to click patterns; interest in specific products and characteristics; visit outcomes; and records of previous visits, where possible, are all examples. As opposed to demographic data, behavioral data does not have to – but sometimes can – be tied to a specific individual. Indeed, "personalization" in this context takes on an expanded meaning: "personas" can be created based on repeated patterns of behavior, and used to guide responses. Responses to an individual customer need not be locked into templates dictated by their age, occupation, income, or even past behavior. Instead, "personas" derived from behavioral data can be dynamic and flexible, reflecting changing moods and objectives.

The development of behavioral analytics, based on this kind of data, has emerged rapidly in the last few years. Just a few years ago, academic approaches focused on creating relatively fixed typologies of consumers to guide action, in an extrapolation of the quantitative marketing models that emerged in the 1980s. Some studies used basic dichotomies, such as browsing – examination of merchandise or service offerings "without a current intent to buy" (Bloch and Richins, 1983) – vs. non-browsing, or recreational vs. informational (e.g., Guiry, et al., 2006). In a more elaborate example, Liu et al. (2012) constructed a typology of four types of "online window shoppers," based on 16 different types of activities. These approaches use limited categories, simple heuristics, and static assumptions. But, as noted, people do not neatly or permanently fit into categories, no matter how sophisticated the classifications are. They are subject to different emotional states, and motivations that vary across time.[16]

An early example of responding to behavioral data in the context of digital commerce was Runa, a company which created a software platform for dynamic responses to site visitors, specifically aimed at increasing conversion rates. In particular, Runa focused on calculating optimal discounts that would convert more visitors to buyers, rather than paying advertising platforms for increased traffic to be driven to a seller's site. The company met success with a free shipping

[16] A pioneering article on these phenomena, in a burgeoning literature in psychology and in behavioral economics, is Tversky and Simonson (1993).

offer program for eBay, using an algorithm for optimal targeting of offers. In 2013, soon after Runa began adding other enterprise customers, it was acquired by Staples, the second-largest online retailer at the time, and the largest seller of office supplies (Guleri, 2013). Since then, of course, the number of firms and the sophistication of tools and applications has exploded. The tools in use are proprietary, but one can get a flavor of the possibilities from research studies (e.g., Behera, et al., 2020). Here, we do not discuss the specifics of any software, but rather the general principles that we view as important, especially in the light of the issues discussed in the previous section.

As noted in the introduction, new and evolving software tools make it possible to work productively with behavioral data. These tools come under the broad heading of Artificial Intelligence (AI). Related terms that are relevant are machine learning, deep learning, data analytics and advanced analytics (AA), although the boundaries between the terms and techniques are not always clear. The boundary between AI and more traditional optimization or analytics depends on the kinds of adjustment or self-correction ("learning") that the software itself is capable of. For example, Biswas et al. (2020, footnote 1) offer this definition: "AI can be defined as the ability of a machine to perform cognitive functions associated with human minds (e.g., perceiving, reasoning, learning, and problem solving). It includes various capabilities, such as machine learning, facial recognition, computer vision, smart robotics, virtual agents, and autonomous vehicles." In some cases, the boundary is difficult to establish: for example, Adobe Systems characterizes its Adobe Target enterprise tool as AI-based, but another perspective is that their use of analytics and optimization falls short of "true" AI. Biswas et al. (2020) handle this fuzziness by typically clubbing "Advanced Analytics" with AI or with machine learning (ML), referring to AA/AI or AA/ML. To some extent, one has to accept that there is no bright line between the different categories.[17]

Turning to the desired goals of employing these new software tools, how can one characterize the process of effectively working with behavioral data? A relatively early attempt to conceptualize a solution was Gartner Research's idea of a Customer Engagement Hub (CEH), viewed as an integrated "system of systems" of software tools. According to a Gartner research

[17] A useful distinction, conceptually and in terms of business implications, is between "strong AI" and "weak AI." See Iansiti and Lakhani (2020) for an insightful discussion.

director, Olive Huang (quoted in Goasduff, 2016), "to offer an end-to-end customer experience across channels and departments, IT leaders must build a CEH. Only a CEH can connect employees across departments, employees with customers, and customers with their peers, while also managing and optimizing personalized customer interaction." Clearly, this is a very expansive vision, one that is encompassing, but difficult to achieve. But the last part of the CEH description, "optimizing personalized customer interaction," captures the central goal of working with behavioral data.

While the CEH remains a conceptual ideal, several other nomenclatures are used to describe current offerings, including "experience optimization platform," "customer engagement platform," or the generic "personalization engine" (Polk et al., 2020). Our own preference is for the term "behavioral data hub," because it focuses on the resource (behavioral data) that is being put to work.[18] The idea of a "hub" here is much more focused than in the case of a CEH. This is a light-touch, adaptable, flexible and modular suite of software that quickly analyzes browsing characteristics (total time on site, speed, page views, and exit intent, in addition to click patterns), interest in specific products and characteristics, visit outcomes, records of previous visits where possible, and more, and provides dynamic, real-time responses, such as discounts, special offers, product recommendations, and so on.

In general terms, these are not new ideas: much of the difference among approaches is in the specifics of implementation. We highlight three such specifics: monitoring, adaptability and data privacy. The last of these, in particular, is an emerging challenge, as described in the previous section.

The ability to monitor outcomes of AI implementations that work with behavioral data is a crucial requirement. Because of the black box nature of how the tools operate (simply because of their complexity and speed, and the overwhelming amount of data that is processed), distilling the results into outcomes that can be meaningfully assessed by experienced humans remains important. Therefore, it is essential to have a customized dashboard that allows for monitoring detailed performance as well as distilling the details to support tactical and strategic goals, and any consequent adjustments in the software that constitutes any behavioral data hub. Most of the

[18] The term "behavioral data hub" is commonly used by Fanplayr, and, as far as we know, was coined within that company.

focus will be on core business metrics such as conversion rates or profit margins, but this kind of monitoring can also be important to avoid unintended social consequences of AI implementations, such as legally-prohibited types of discrimination.[19] Such characteristics may seem obvious, but consider the following observations by an industry analyst with respect to one such effort, "While Dynamic Yield made improvements to measurement and reporting – especially for advanced users through a performance dashboard – client references cited issues with data transparency, particularly with visibility into how data is used and with verifying specific recommendations. They also reported bugs and difficulties with data queries for reporting across multiple campaigns."[20]

It is also worth emphasizing that adjustments based on human intervention have to be judicious. One cannot fall into the opposite trap of having to rely on a team of people to be constantly coding. This is one part of our second desideratum for a behavioral data hub, adaptability. In the case of another well-known implementation, Adobe Target (an enterprise tool offered by Adobe Systems), the software offers many potential capabilities, but is complex to manage, making it cumbersome and expensive in practice. This example illustrates the complex challenges of working effectively with behavioral data: not all "AI" is equal.[21]

The second part of adaptability has to do with flexibility and scope. For a behavioral data hub to be useful, it cannot be a single-action solution, such as for email retargeting, testing, or discount offers, where decisions on whether an action is taken or not have to be pre-committed. In that sense, technology and technological potential have moved beyond the kind of earlier, narrower approaches illustrated by cases such as Runa. Indeed, the ability of a well-designed solution to

[19] A well-known example is Amazon's choice of neighborhoods where it rolled out one-day shipping – their algorithm effectively "redlined" areas with higher African American populations. See, for example, Bruckner (2018), which discusses this issue in the context of credit markets, in particular.
[20] See Polk, et al. (2020) for the complete discussion of this and other examples. It is important to note that this is a successful effort, since the company was acquired by McDonalds. In the case of an in-house operation for a large corporation, customized dashboards at this level may seem less important relative to the parent company's overall monitoring requirements and management goals. However, behavioral data analytics efforts that are meant to serve a range of clients cannot neglect this issue.
[21] Again, we rely on Polk et al. (2020): "Adobe Target lacks the native ability to collect user feedback, requiring custom partner integrations for customer survey design, execution and data collection. Client references also rated the company below average for its ability to drive CX [customer experience] personalization through its personalization engine." As noted earlier, our perspective is that their use of analytics and optimization falls short of "true" AI, even "weak AI."

"read" intent from digital behavioral data and adjust responses from a flexible set of options ought to be a necessary feature of the use of AI in these contexts.[22]

The third and final aspect of a behavioral data hub for the digital economy of the next decade is how it handles issues of privacy. In the previous section, we reviewed how a combination of concerns over market power aspects of digital agglomerations of personal data and over the erosion of general norms of personal privacy has led to increased regulation of data collection and use. Clearly, this increases the importance of behavioral data collected in first-party contexts, such as seller or service provider websites, versus the indiscriminate personal data amassed by some of the digital giants. Specifically, as the controls on user privacy gets tighter, more and more of the online browsing activity will be anonymous. Where there is anonymous activity, unless and until a user 'logs in' to identify themselves, there is nothing that can be done to provide a more personalized experience, unless one uses behavioral data to infer their intent.[23]

However, the technical challenges raised, in using behavioral data effectively are nontrivial. To illustrate, changes in regulation, technology and use cases, and the proliferation of tracking services have caused web browser providers to step up their drive against overly broad tracking. Every browser vendor has taken a different approach in their implementation of tracking prevention. One thing consistent between them is that they are all making it increasingly difficult to track a user's journey across sites. In this situation, intelligent analysis of behavioral data that robustly and consistently classifies users (even without demographic identifiers), and segments and responds to them appropriately, is an important quality for a behavioral data hub.[24] Two of the authors' direct experience with implementing this approach disproves, we argue, what seems to be the general assumption that data is only useful when collected across sites and in a broad browsing context. Indeed, the more a behavioral data hub is used, the more useful it becomes.

---

[22] As a simple example, an existing bank customer seeking to make an appointment could be offered a new savings product, based on the behavioral data generated during the process combined with historical data.

[23] Multi-user devices (especially mobile) also necessitate the use of behavioral data to understand intent.

[24] Of course, a known customer, who has logged into a site, is still identifiable – their deliberate action has provided consent for this identification. On the other hand, behavioral data allows for greater flexibility than "personalization," if the latter term is interpreted as being tied to fixed personal or demographic characteristics. Instead, a focus on behavior allows for the multiple layers of people's identities to be recognized and responded to. This difference in focus relates to concerns about the use of people's demographic data in ways that do not respect privacy. The privacy issue is broader, of course, because it addresses the use of behavioral data as well, when that is done in unsanctioned ways, such as leveraging online social interactions into commercial activities. Again, the Chinese model of digital commerce represents a different attitude toward these issues.

To round out this discussion, we note that the idea of a focused behavioral data hub does not exclude broader developments: digital commerce obviously requires a complex of software components, all working together. What we emphasize here is a focus on what can be termed the "digital last mile," where potential and actual customers and clients engage intensively with sellers and service providers.[25] To put this in perspective, spending on this stage of individual digital journeys is as little as one percent of spending on driving traffic to websites. This statistic reflects the unbalanced pattern of development of digital commerce, as well as the market power that goes with that pattern.

Implementation: Illustration and Implications

In this section, we illustrate more specifically how behavioral data can be used effectively, outlining a specific implementation for doing so: "Method and System for Segmentation as a Service."[26] The central approach involves delivery of instrumentation code and "segmentation processing code" to a user's machine. This enables intelligent and actionable tracking of a user session on a website.

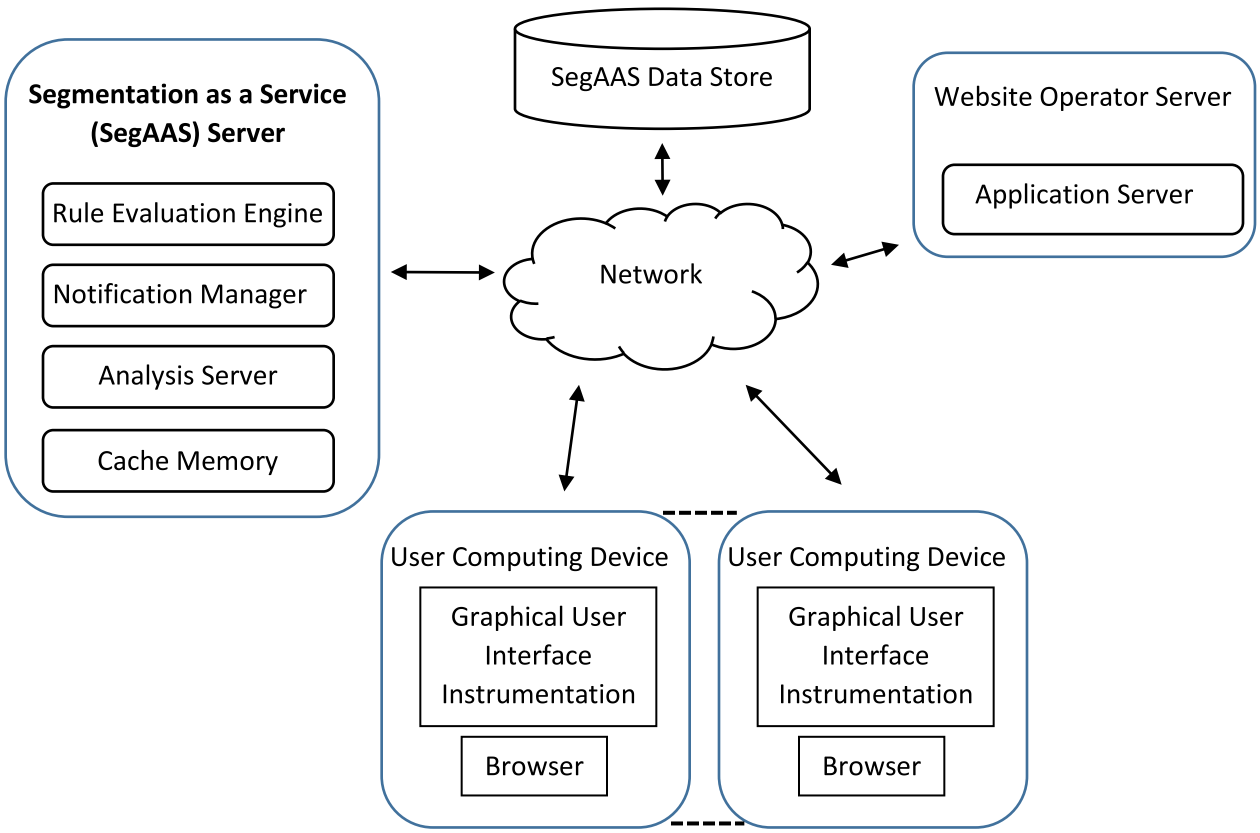
More specifically, data from the user session is received and processed, and callback function code is invoked, configured to run on the user's machine. In turn, this code performs actions that can be specified by the website's operator, as triggered by delivery of segment codes. The process involves dynamic feedback loops. Readings from the user session are received continually, and that data enables automatic tracking of the progress of the session. Furthermore, immediate processing of this session data generates segment codes that characterize results of the analysis in an actionable way. This updating is based on rules that can be specified by the website operator, based on appropriate business objectives (e.g., increasing the conversion rate, or offering product recommendations). Delivery of updated lists of the segment codes can be repeated, based on analyzing the progress and timing delivery of unsolicited actions, which may be directed to retaining the visitor and extending the user session with the website, or to successfully completing an action (e.g., making an appointment) or transaction.

[25] Again, this focus does not rule out a "system of systems" or, ideas that are emerging in fintech, of "open banking" or banking "ecosystems."

[26] This discussion is based on a patent filing.

As discussed in previous sections, the technical approach just outlined uses relevant behavioral data in an efficient manner, and does not have to rely on large amounts of demographic and peripheral behavioral data accumulated from across multiple websites. It represents a significantly less costly and more effective approach to targeting, segmentation and personalization. As with any such software implementation, there are numerous technical details, but the central features of the implementation are illustrated in Figure 1.

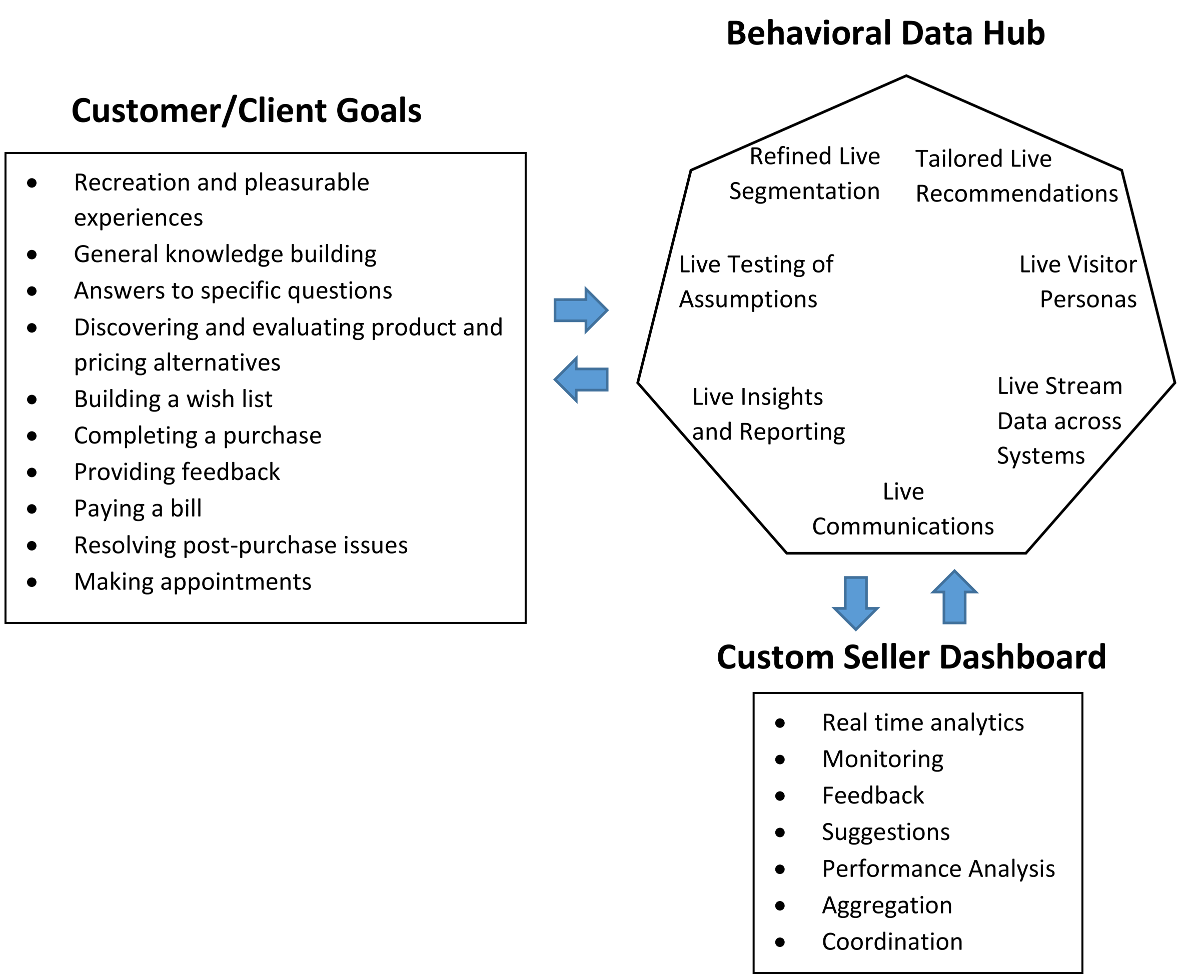
Figure 1: Components of Segmentation as a Service



One important feature to note is that the implementation can operate across multiple user devices, and multiple kinds of devices (e.g., laptops and smartphones), operating systems, and user interfaces (browsers and apps). This is obviously a necessity in the modern computing access environment, and is relatively standard. Where this approach innovates is in the ability to effectively process behavioral data in a manner that is simultaneously more privacy-respecting and more supportive of business objectives. Of course, the figure is simplified, but it illustrates the components of a system that realistically tackles the challenges identified in earlier sections of the paper.

The illustration of the privacy-respecting, "last mile" personalization approach just presented also has some operational and managerial implications, which we outline next. Typical software solutions that tackle the core issue of effective personalization on a seller or service-provider's website offer bundled solutions of simple segmentation combined with targeting capabilities, but mostly fail to respond adequately to the user's intent. These limits are often addressed by tweaking the website regularly using in-house engineering resources, or pushing that task on to hosting software that develops complexities which are hard to manage and reduce transparency and speed of innovation.

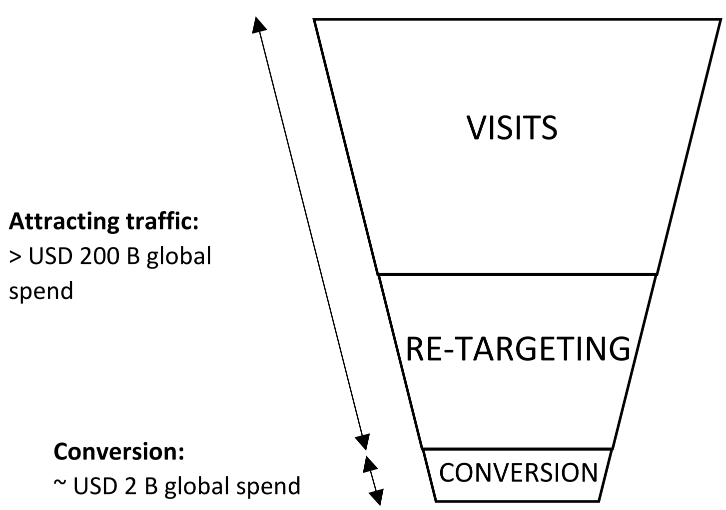
Figure 2: Managerial Implications of Segmentation as a Service



Websites also face a common problem of mixing the roles of a technical resource specialist who provides development skills, a design resource who provides user interface (UX) skills, and a marketing resource who provides an understanding of the market and customers. In many typical software solutions to "last mile" personalization, segmentation and targeting are rolled into a single product, and these roles can get merged. This makes it difficult for a marketing person to manage the segments and design widgets, including a possible need to build small scripts on the website. The approach we have outlined solves this problem through an effective use of AI, its integration into the seller's overall information systems. A customizable dashboard can then provide managers with flexible monitoring and adjustment tools. The managerial implications of all of this are illustrated in Figure 2, where the relevant AI software suite is described as a "behavioral data hub."

To make these ideas more concrete, we focus on two bottom-line outcomes: conversion rate and average order value (AOV). Clearly, increasing either or both of these will increase revenue. Because of the uneven development of the commercial aspects of digital networks, as discussed earlier in this paper, the resources devoted to achieving conversion of visitors to e-commerce websites are dwarfed by the amounts spent on driving traffic to those sites, by a factor of over 100, according to our estimates. This is illustrated with a "conversion funnel," in Figure 3.

Figure 3: The Conversion Funnel



The bottom-line impacts of intelligent analysis of visitor behavior are illustrated in Table 1. The approach discussed earlier in the paper, and illustrated in Figures 1 and 2, tracks and analyzes visitor intent in real time, for example by identifying gaps in browsing activity that might indicate temporary distraction, and responding appropriately. Clearly, using this kind of behavioral data is more privacy-respecting than the "traditional" use of demographic data collected and aggregated from multiple sources. The data are based on standard A/B testing, and are indicative of how this approach can yield tangible benefits for digital commerce. In both cases, there are non-trivial effects on AOVs and conversion rates. An important, but less obvious implication of this data (e.g., the different conversion rates for the two types of sellers) is the importance of customization. An approach that focuses on the digital last mile and real-time behavioral data is better suited for this than the usual indiscriminate agglomeration of personal data from varied contexts, such as generalized search or social networking.

Table 1: Behavioral Data – Bottom-Line Impacts

	Travel Accessories Site (US)			Fashion Clothes Site (Japan)		
	Control Group	Target Group	Percent Change*	Control Group	Target Group	Percent Change*
Visitors	5,539	5,600	n/a	14,304	14,443	n/a
Total Net Revenue	$58,787	$73,872	24.3	¥1.899M	¥2.445M	27.5
Number of Orders	737	872	17.0	388	441	22.8
AOV	$79.74	$84.72	6.2	¥4,894	¥5,082	3.9
Conversion Rate	13.31	15.57	17.0	2.71	3.33	22.9

*Percentage changes are adjusted for differences in control and target group sizes.

The sizable revenue gains in the examples also deserve comment. Despite the differences in the details of the nature of the products, onsite behavior, and factors such as AOV and conversion rates, gains of this magnitude consistently occur in other experiments as well. The question then arises as to why this situation persists. Based on the arguments in this paper, part of the explanation is in the unbalanced development of certain aspects of the commercial Internet, as discussed in this paper and encapsulated in Figure 3. The other part of the explanation is in the relatively recent development of software that both uses AI appropriately, and allows for light-touch integration with existing software systems. As the potential of this approach becomes more widely recognized and adopted, economies of scale will kick in, and accelerate the transition from the current equilibrium.[27] As we have discussed, regulatory developments and consumer concerns with respect to privacy will also be factors in this transition.

Conclusion

In this paper, we have reviewed emerging issues with respect to market power and user privacy in digital commerce. These issues have surfaced because of novel sources of economies of scale in digital networks, including the advantages of amassing data on users of these networks. They reflect an unbalanced development of digital commerce. New restrictions on using third-party data indiscriminately, along with advances in various forms of AI, machine learning, and related software tools, open up possibilities for new directions of evolution.

In particular, the new software tools create possibilities for sellers and service providers to regain some control over their digital strategies, relying less on the pervasive presence of a few technology giants. Specifically, behavioral data – both live and historical – that is collected in the digital last mile in well-defined first-party contexts allows for a range of more specialized sellers and service providers, in sectors such as retailing, hospitality and travel, financial services, and telecom services, to engage effectively with customers, clients, or various other kinds of visitors to their websites (or their apps, for existing users). By reducing the reliance on static demographic data, this also opens up the possibility of more flexible and expansive implementations of "personalization."

We termed the mechanism for implementing this approach a "behavioral data hub," and discussed monitoring, adaptability and privacy management as desirable characteristics of such an implementation. We also distinguished the concept of a behavioral data hub from more sweeping ideas of integrated "systems of systems" of software tools, which encompass the entire digital footprint of an enterprise. These idealized visions can be difficult and costly to achieve, whereas the behavioral data hub, implemented as we outlined here, is practical and cost-effective, and has important managerial implications in terms of how technology can be deployed productively by business-side decision makers. This focus on the digital last mile is feasible, with tangible business benefits, as well as an opportunity to correct some of the

[27] We would also suggest that the specific application of AI discussed in this paper is an important example of the general argument presented in Iansiti and Lakhani (2020).

imbalances in market power and data privacy that have arisen in recent years with respect to the functioning of digital networks and digital commerce.

References

Bamberger KA, Lobel O (2017) Platform market power. *Berkeley Tech. L.J.* 32(3):1051-1092

Bariso J (2020) Tim Cook May Have Just Ended Facebook. *Inc.* January 28, https://www.inc.com/justin-bariso/tim-cook-may-have-just-ended-facebook.html

Behera RK, Gunasekaran A, Gupta S, Kamboj S, Bala PK (2020) Personalized Digital Marketing Recommender Engine. *Jnl of Retailing and Consumer Services*, 53, March, 101799, https://doi.org/10.1016/j.jretconser.2019.03.026

Biswas S, Carson B, Chung V, Singh S, Thomas R (2020) *AI-bank of the future: Can banks meet the AI challenge?* McKinsey & Company, Global Banking & Securities, September

Bloch PH, Richens ML (1983) Shopping Without Purchase: An Investigation of Consumer Browsing Behavior, in *NA - Advances in Consumer Research Volume 10,* eds. Bagozzi RP, Tybout AM. Ann Arbor, MI: Association for Consumer Research, 389-393

Bruckner M (2018) The promise and perils of algorithmic lenders' use of big data. *Chicago-Kent Law Review* 93(1): 3-60

De Groot J (2019) What is the General Data Protection Regulation? Understanding & Complying with GDPR Requirements in 2019, *DataInsider: Digital Guardian's Blog*, July 15, https://digitalguardian.com/blog/what-gdpr-general-data-protection-regulation-understanding-and-complying-gdpr-data-protection

Evans D, Schmalensee R (2007) The industrial organization of markets with two-sided platforms. *Compet Policy Int* 3:151–179

Gaus PR (2017). Only the good regulations die young: Recognizing the consumer benefits of the FCC's now-defunct privacy regulations. *Minn Jnl Law, Sci & Tech* 18(2):713-744

Goasduff, L (2016) Ten Steps to Plan a Next-Generation Customer Engagement Hub. Gartner Inc., May 10, available at http://www.gartner.com/smarterwithgartner/ten-steps-to-plan-a-next-generation-customer-engagement-hub/

Graef I (2015) Market Definition and Market Power in Data: The Case of Online Platforms'. *World Competition* 38 (4): 473–506

Guiry M, Magi, AW, Lutz RJ (2006) Defining and Measuring Recreational Shopper Identity, *Jnl Acad of Marketing Sci*, 34 (1): 74-83

Guleri T (2013) Staples Hits the Easy Button in Big Data and Acquires RUNA. October 3, *Sierra Ventures*, http://www.sierraventures.com/2013/10/staples-hits-easy-button-big-data-acquires-runa-2/

Haucap J, Heimeshoff U (2014) Google, Facebook, Amazon, eBay: Is the Internet driving competition or market monopolization? *Int Econ Econ Policy* 11:49–61

Iansiti M, Lakhani K (2020) Competing in the Age of AI: How machine intelligence changes the rules of business. *Harvard Business Review*. January-February: 61-67

Katz M, Shapiro C (1985) Network externalities, competition, and compatibility. *Am Econ Rev* 75:424–440

Liu F, Wang R, Zhang P, Zuo M (2012) A Typology of Online Window Shopping Consumers. *PACIS 2012 Proceedings* Paper 128, http://aisel.aisnet.org/pacis2012/128

Maple C (2017) Security and privacy in the internet of things. *Journal of Cyber Policy*, 2 (2): 155-184

Nuccio M, Guerzoni M (2019) Big data: Hell or heaven? Digital platforms and market power in the data-driven economy. *Competition & Change* 23(3):312–328

Polk J, Tassin C, McNellis J (2020) Magic Quadrant for Personalization Engines. *Gartner Report Reprint*, 13 July, ID G00439919

Prado, TS (2020) Assessing the Market Power of Digital Platforms, Working paper, Michigan State University, available at: https://ssrn.com/abstract=3747793

Rice JC, Sussan F (2016) Digital privacy: A conceptual framework for business. *J of Payments Strategy & Systems* 10(3): 260-266

Rochet JC, Tirole J (2003) Platform competition in two-sided markets. *J Eur Econ Assoc* 1:990–1029

Rohlfs J (1974) A theory of interdependent demand for a communications service. *Bell J Econ Manag Sci* 5:16–37

Sarathy R, Robertson CJ (2003) Strategic and Ethical Considerations in Managing Digital Privacy. *Journal of Business Ethics* 46: 111–126

Singh N (2020) Digital Commerce in the Post-COVID Network Age, Fanplayr Whitepaper, October, available at https://fanplayr.com/files/Digital Commerce Post-COVID - October 2020.pdf

Sisk EP (2016) Technical Difficulties: Protecting Privacy Rights in the Digital Age. *New Eng J on Crim & Civ Confinement* 42(1):101-120

The Economist (2021) The next big thing in retail comes with Chinese characteristics. January 2 https://www.economist.com/business/2021/01/02/the-next-big-thing-in-retail-comes-with-chinese-characteristics

Tversky A, Simonson I (1993) Context-Dependent Preferences. *Management Science*. 39 (1): 1179-1189

Weber RH (2015) Internet of things: Privacy issues revisited. *Computer Law & Security Rev* 31: 618–627