

This PDF includes a chapter from the following book:

Assetization

Turning Things into Assets in Technoscientific Capitalism

© 2020 Massachusetts Institute of Technology

License Terms:

Made available under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International Public License

<https://creativecommons.org/licenses/by-nc-nd/4.0/>

OA Funding Provided By:

The open access edition of this book was made possible by generous funding from Arcadia—a charitable fund of Lisbet Rausing and Peter Baldwin.

The title-level DOI for this work is:

[doi:10.7551/mitpress/12075.001.0001](https://doi.org/10.7551/mitpress/12075.001.0001)

3 Datasets: Assetizing and Marketizing Personal Data

Thomas Beauvisage and Kevin Mellet

Introduction

In a 2011 report entitled “Personal Data: The Emergence of a New Asset Class,” the World Economic Forum (WEF) stated that “personal data is generating a new wave of opportunity for economic and societal value creation. ...As some put it, personal data will be the new ‘oil’—a valuable resource of the 21st century. It will emerge as a new asset class touching all aspects of society” (Schwab et al. 2011, 5). Starting from this observation, the report called for the establishment of a new data infrastructure, which would simultaneously give individuals high control over their personal data and facilitate the flow and exchange of data so far retained in isolated silos among individuals, companies, and public institutions. The report predicted that personal data would constitute a new form of currency for individuals—one that they would be free to manage themselves: “In practical terms, a person’s data would be equivalent to their ‘money.’ It would reside in an account where it would be controlled, managed, exchanged and accounted for just like personal banking services operate today” (Schwab et al. 2011, 10).

Admittedly, the economic importance of personal data is not new. In the 1970s (to stick to contemporary times), private players developed database management software, and some of them, which later became known as data brokers, established the basis for the trading of lists (of addresses, contacts, profiles, credit reports, etc.). It was during this period that an academic literature on data and privacy (Posner 1981) and modern legislation on the protection of personal data developed (Fair Credit Reporting Act, USA 1970; Loi Informatique et Libertés, France 1978). That said, the recent and combined accelerations in the areas of networks, storage, and

computing, and the multiplication of sensors that keep a record of various traces have led to a renewed interest in data, sometimes labeled as the “big data revolution” (Kitchin 2014). As pointed out by Schneier, “in the normal course of their operations, computers continuously document what they’re doing” (2015, 15). Big data systematizes the apprehension of data and represents a valuable resource requiring attention and investment.

The big data revolution raises a host of questions. How to make money, and more generally extract value, from personal data? How concretely are personal data transformed into economic assets that generate revenue streams and value? How does this new asset class emerge? The WEF provides a typically neoliberal answer to this question, but it is not the only possible one. Arguing for a supposed “exceptionalism” of big data (Nissenbaum 2017), large technology companies and a multitude of start-ups collect huge amounts of user data and put together teams of data scientists to analyze and extract value from these data. Traditional companies have gone through an assetization movement in recent years as well. They have set up data management entities headed by chief data officers and assigned them the task of identifying valuable first-party data within their complex information systems in order to derive revenue streams from them.

Building on a decade-long inquiry into the online marketing and advertising industry, this chapter analyzes the process by which consumer data have come to be considered as valuable, and sometimes tradable, assets. In the introductory chapter, Birch and Muniesa defined assets as “something that can be owned or controlled, traded, and capitalized as a revenue stream, often involving the valuation of discounted future earnings in the present.” This precise characterization of assets appears as a distant horizon when one is interested in personal data. Indeed, the technical, legal, and economic contours of personal data appear hazy and fluctuating. Moreover, contrary to what is observed in finance, there is no set of standards and instruments for measuring, assessing, and discounting future revenue streams from personal data. Finally, market-based trading of personal data by data brokers and similar actors is only a small part of a much larger transformation that implies the production, activation, and flow of data, at a large scale. As underlined by Muniesa et al., anthropology and sociology typically focus on “commercial valuation—things being valued as they are bought and sold, and hence valued as commodities in the market” (Muniesa et al. 2017, 13), but pay little attention to the process of capitalization.

We argue that the ability to capitalize personal data in the present is the result of a versatile and uncertain process of assetization. This chapter seeks to unpack this process and to identify and qualify various data activation regimes and the resulting statuses of data as assets, or “datasets” as we name them.

The online marketing and advertising industries are a striking example of the dynamics of data assetization, a movement also observed in areas such as healthcare (Ebeling 2016), insurance (McFall 2014), and government assistance (Eubanks 2018). They provide a notable illustration of something that economic institutions and mainstream economic thought have considered explicitly as an asset, while letting firms and market intermediaries perform this claim. Our examination of how economic players have succeeded or failed in their attempts to assetize personal data unpacks the temporal and entrepreneurial dynamics of assetization, their material and discursive devices, and highlights the plurality of asset forms that emerge from this dynamic.

The chapter is structured as follows. The first section presents a brief survey of the literature. Next, we analyze the repeated and unsuccessful attempts to create a consumer-to-business (C2B) market for personal data, and the underlying conception of personal data marketization as a trade-off between privacy and the benefits of tailored services. This failure can be considered as a performativity failure of economic models and related experiments. In the third section, we investigate the actual markets for personal data, a business-to-business (B2B) activity strongly linked with marketing and advertising services. We show that personal data can be commodified and traded, especially by data brokers, but as resources adding value to advertising products in the form of contacts, segments, or attributes. We finally argue that the spreading of tracking technologies and data management platforms (DMPs) inside private companies contributes to transform datasets into “datasets,” that can either compete or articulate with third-party, commodified data.

Putting Personal Data on the Research Agenda

Within the field of marketing and advertising there is an emerging literature seeking to open the black box of data marketing by describing its intermediaries, products, processes, and organizing principles. In his book

The Daily You, Turow (2012) investigates the development of the online advertising market. He presents an entire industry engaged in a process of splitting Internet users into “targets and waste.” Turow argues that a process of personalization is under way, based on the creation of extremely detailed databases on the characteristics and tastes of Internet users. At the end of his book, he focuses on data brokers, a set of emerging actors specialized in the collection, aggregation, processing, and selling of personal data, which tend to gain importance in the ecosystem of online advertising. Data brokers, like Acxiom or Experian, are shrouded in secrecy: their activities were revealed to the public by an investigation conducted by the Federal Trade Commission (2014). The corresponding FTC report provides a valuable description of this industry: its data acquisition methods, its products (segments and scores for marketing, risk mitigation for finance, people search), and its customers. This report serves as a primary source for works that seek to account for the breadth and variety of forms of tracking and surveillance put in place by private companies in the 2010s (Christl and Spiekermann 2016). Conversely, some authors notice how increasing concern about privacy has been seized as a business opportunity by entrepreneurs (Milyaeva and Neyland 2016).

Other research examines the forms of stranglehold and appropriation at work on the part of this industry and questions the moral and legal legitimacy of this foreclosure. Ebeling (2016) is interested in how ordinary life events become objects of capitalization by data brokers. In an ethnographic inquiry into her “marketing baby,” she describes the logics of capture implemented by data brokers and how they are coupled with a form of appropriation. For her, the transformation of private data into assets is based on an ownership claim. “Brokers argue that by ‘adding value’ to data through analyzing it, processing it, de-identifying it, and creating new instruments or data products out of ‘raw’ data, brokers firm up their ownership claims. A common refrain in the industry is that data are the ‘new oil’ and brokers are the ‘processors’ that refine it and make it into ‘products’” (2016, 44).

Nissenbaum (2017) is interested in the legal arguments put forward by such big data promoters, especially those engaged in an ideological and political battle to eliminate the regulations covering the collection of data. According to them, it is the use of data, not its collection, that must be regulated, a stance Nissenbaum labels “big data exceptionalism.” In other words, personal data constitute a store of (economic, social, etc.) future value, but

the uncertainty about the nature and the modalities of extraction of this value are framed such that the recording of traces and constitution of giant databases should not be hindered. Following a complementary perspective, Fourcade and Healy (2017) point out that organizations follow an institutional data imperative to collect as much data as possible, which results in the constitution or reinforcement of a new form of übercapital built from these data, in particular in the form of rankings and individual scores.

These works provide a solid basis for analyzing how personal data are constructed as economic assets in industrial and market processes, although they do not focus on the devices put in place by corporations to transform data into assets. Rather, capitalization is often reduced to a binary sequence of operations of capture and marketization, and carried out by data brokers and other market intermediaries only.

Attempts to Marketize Personal Data

Around 2010, personal data appeared to be a new asset for a large range of economic players, strongly linked with the digital economy and often described as the “new oil” in a future fully digitized world. But the potential value of this new resource was still to be determined. To extract a share of this new value, a part of the digital economy considered personal data as a commodity, a standard elementary tradable good like oil, close to a currency. Such perspectives on personal data as tradable assets was reinforced by the forecasts of neoliberal institutions such as the World Economic Forum, which imagined a direct conversion of personal data into revenue streams, as we mentioned above.

“New Assets” as Possible Products and Markets in the Future

The WEF’s initiative is not isolated. It took place in a doubly favorable context. On the one hand, in the digital economy, entrepreneurs glimpsed an opportunity to build business models around the marketing by individuals of their own data. On the other hand, within the civil society, individual rights groups such as the Electronic Frontier Foundation have supported initiatives aimed at empowering individuals against the gluttony of large digital service corporations. In this context, it seemed that there was a genuine prospect of a market for personal data in which Internet users would make money from selling information about themselves. In order to sketch

the outlines of such a system, academic research and consumer surveys led by private companies attempted to appraise the market worth of personal data. These economic experiments all take the same approach, immersing Internet users in fictitious situations in which they have to weigh up the cost of disclosing information about themselves (age, income, Internet history) against a financial benefit (discount or revenue). The value of data is thus defined as the price—or its equivalent in service terms—at which individuals would agree to transfer specific pieces of private information. The aim of these initial valuations was to identify the data whose protection individuals valued most highly, as well as the third parties to whom they would consider selling them and at what price. It was also a way to assess the cost of privacy (see Table 3.1).

Table 3.1

Summary of studies on the worth of personal data item for individuals

Source	Average valuations
Huberman et al. (2005)	Age = \$57 Weight = \$74
Danezis et al. (2005)	Location ~ £27
Beresford et al. (2012)	Favorite color and year of birth = €1
Carrascal et al. (2013)	“Offline” information (age, address, economic status) ~ €25 Browsing history = €7 Interactions on social networks = €12 Search history = €2 Shopping = €5
Staiano et al. (2014)	Geolocation = €17 to €588 Communications = €3.40 to 51
Havas Média (2014) ^a	30% sell in the top band, i.e., “More than €500”
Orange (2014) ^b	Each piece of information ~ €15 (name, mobile number, children’s ages, income, purchase history, contacts, etc.) Average total value = €170

^aHavas Media. 2014. Les Français et leurs données personnelles, quelle place pour les marques? Corporate report.

^bOrange. 2014. The Future of Digital Trust. A European Study on the Nature of Consumer Trust and Personal Data. Corporate report, <http://www.orange.com/content/download/25973/582245/version/2/file/Report+-+My+Data+Value+-+Orange+Future+of+Digital+Trust+-+FINAL.pdf>.

The study conducted by Carrascal et al. (2013) is a good example. The experiment was based on a plug-in developed for web browsers, which, in real time, asked a panel of participants (n=168) at what price they would be willing to sell information about their behavior and practices on the web. “Offline” information (age, address, economic status) had a high price, at around €25. The average price asked by participants for a permanent transfer of their browsing history, meanwhile, was €7, with variations depending on the information (€12 for interactions on social networks; €15.50 for visits to financial sites; €2 for search history; €5 for shopping).

Unsuccessful Attempts to Create a C2B Market for Personal Data

Some start-up companies explored this path, banking on the idea that digital players would respect the need to seek users’ consent to provide their data and on the notion of attaching proprietary rights to personal data, thus turning it into an asset. We have identified five such players, in the United States (Personal, Datacoup, Personal Black Box), the United Kingdom (Handshake), and France (Yes Profile). These five companies were established between 2009 and 2013.

These five start-ups all used the same pitch, which can be summed up as follows: the business model of Internet giants is based on the unfair seizure of private information from which they make enormous profits. Instead, these start-ups would act as personal information brokers on behalf of their subscribers toward the advertising industry, redirecting profits and making users’ consent a central tenet. The model adopted by these companies involved asking Internet users to entrust them with data from social networks, as well as entering a considerable amount of additional information about themselves, particularly about their purchase intents, or even taking part in focus groups and testing products. The potential gains these start-ups promised to web users were in some cases highly exaggerated: Handshake suggested the revenue stream could reach £15,000 (€19,000) per participant per year; Yes Profile estimated potential monthly income at €60 to €100, giving an annual total of up to €1,200; and Personal Black Box gave a range from \$50 to \$500 (€45 and €450). Datacoup, meanwhile, was more cautious, with its site advancing the figure of €8 a month, or around a hundred euros a year.

This business model never took off. Ultimately, it came up against one major stumbling block: the capture and sale of web users’ traces and

personal information was already a widespread practice, and it did not stop while users subscribed to their services. The legal framework is such that advertisers do not need to turn to Internet users to access information that they already systematically obtain directly through their websites and applications (data from CRM databases, behavioral data from their website), or indirectly through advertising and marketing players. The founder of Personal, Shane Green, observed in 2013 that the market for personal data “does not exist right now, because consumers are not in on the game.” A counterargument might be that it is precisely because consumers are not in on the game that the market *does* exist today (discussed below). In other words, companies like Yes Profile, which promised individuals that they can “own their profile again,” are unable to keep this promise: signing up on Yes Profile website and providing one’s personal information did nothing to stop the tracking performed by other players (advertising networks, ad exchanges, data brokers, etc.). The profiling performed by Yes Profile only adds to the multiple tracking mechanisms individuals are exposed to, it does not replace them. Consequently, the C2B market for personal data has never really taken off, and the start-ups that have tried to create it are either doing negligible business or have changed their core business, abandoning personal data monetization. Personal fits into the latter category, now selling a collaborative software solution. As for Yes Profile, its web page was inactive in November 2019 and the company seemed to have left the business.

In addition, even if Internet users were to take back control of their data and their traces and try to make money out of them, the self-valuation of personal data by Internet users would come up against another major constraint: on existing markets, individual level information is not worth much, and this effective market price probably does not justify individuals going to the effort of putting them up for sale. Calculations from two sources based on the rates used on the advertising market provide empirical support for this observation. The *Financial Times* was the first, in June 2013 (Steel et al. 2013) to assess the value of typical pieces of personal data (sociodemographic information, assets, leisure pursuits, and consumption patterns) based on list prices from data brokers in the United States. Summing the value of each piece of information, the average total value was running at around 20 cents. A second empirical study by Olejnik et al. (2014) reports purchase prices in online advertising auction systems (known as real-time bidding or RTB). The authors observe that browsing

history itself is only worth around \$0.0005 for an advertising impression; when applying that price to the average number of sites visited and advertisements displayed per site, the authors estimate the volume of business generated at \$0.18 per user per month. Web users' potential earning on this market would be around \$0.432 per year.

The failure of the World Economic Forum's forecast of personal data as an "equivalent to 'money' ...controlled, managed, exchanged and accounted for just like personal banking services" (Schwab et al. 2011, 10) for individuals can be seen as performative failure of neoliberal ideology. Although personal data can indeed be considered as an asset and a tradable object, the mistake was to consider that its potential value would be unleashed through a C2B marketization process. The idea of personal data as liquidity is also the avatar of the attempt to gauge the potential value of something without being able to dig it into markets. In that sense, the failure of a C2B market for personal data also signals that personal data cannot be considered as a good as-is, but need to be constituted as products to match operational needs and take a specific place in market architectures. In fact, the market for personal data as a product has already existed for decades in the marketing area: a B2B market for purchase intentions, qualified profiles, and lists of individuals, framed around the needs of advertisers and a highly automated industrial supply chain.

The Assetization of Personal Data in the Marketing Industry

The actual market for personal data has its roots before the development of the Internet. It is a B2B affair, with private and public organizations as its main customers, and long-established providers, particularly in the marketing and advertising fields. The activities of marketing and advertising professionals fall into two distinct categories: direct marketing and media advertising. In both cases, the specialists operating on these markets can gain a lot of efficiency in their operations if they have information on how to address their targets. They did not wait for the big data revolution to introduce tools for aggregating, processing, and cross-referencing personal data, turning them into a tradable good between companies. However, the transformations in progress should not be underestimated: we are witnessing large-scale and highly automated industrial engineering of the capture, cross-referencing and use, of individual data. The driving force behind it is

the transformation of the central product traded on the advertising market—namely, the audience (see also Turow 2012). In this section, we show that personal data are not traded by data brokers and advertising networks as a commodity, a generic material that would fit a large array of purposes, but rather as refined input tailored to fit the needs of the advertising industry and its products. Personal data are traded in the form of (addressed) contact lists and audience segments, in order to serve as an upgrading and valuing tool for the main products that are traded by advertising and direct marketing firms: ad impressions and direct emails. In that sense, personal data do not constitute a commodified good but rather an asset upon which value can be built, a lever for marginal gains in an already well-equipped market.

Direct Marketing: From List Sellers to Data Brokers

Companies' direct marketing activities target consumers through direct means, as opposed to indirect means (the media). The traditional methods are direct mailing and cold-calling. Over the past twenty years or so, these have been joined by direct emailing. These activities have led to the emergence of an address rental market. This market serves a dual purpose for advertisers, providing them firstly with lists of addresses, preferably up-to-date and with names attached, and secondly with addresses that are qualified—in other words, paired with information that indicates whether or not the concerned individuals are potential purchasers of their products.

The history of address rental, which really took off when companies began to computerize in the 1970s, provides an insight into the place of personal data in today's business practices. Here are the major elements of the story, as told by one of its historic players, French firm ITL: "The early sixties saw many companies specialize in supplying the addresses of people from certain socio-professional categories—doctors, dentists, engineers, students at top universities—collectively referred to, rather eloquently, as high-purchasing-power individuals."¹ In the late seventies, the computerization of small and medium sized enterprises (SMEs) boosted this market through over-the-counter transactions between companies, particularly mail order players, who took the risk of making the names and addresses of their customers available to other companies in exchange for an identical volume of addresses or payment based on a price per thousand.

How did we move from a local and almost artisanal activity in the 1960s to international players building consolidated databases with thousands of

variables covering large swaths of the population? Information technology was the key to the change, both for list suppliers and data brokers and for their customers. For the latter, digitization meant not only computerized CRM databases, used to calculate lead scores but also the widespread implementation of loyalty programs, after airplane companies implemented frequent flyers programs in the 1980s (Araujo and Kjellberg 2016) and participated in the spread of this new market device in the retail industry (Pridmore 2008; Coll 2013). Companies thus accumulate knowledge of their customers that includes transactional elements. They can use these elements firstly to get to know their customers better and adapt their pitch and offers accordingly, and secondly to identify prospective customers and customer catchment areas.

Industrial concentration was the main factor leading to the emergence of data brokers—such as Acxiom, Experian, and Epsilon—who built mega databases accumulating as many variables as possible on consumers. The history of the data brokerage companies, as partially unveiled by data brokers themselves (Watson 2013) or academic works (Goss 1995; Hoofnagle 2003; Bouk 2017), shows how today's large players are the result of small companies initially specializing in some particular calculation on individual data: credit scores (Experian, Equifax), political marketing (Acxiom), loyalty programs infrastructure (Alliance Data). As long as these companies needed to have as much customer information as possible, large movements of consolidation led to today's large multipurpose data brokers. Although data brokers are fueled with individual information from various sources, they do not sell personal data themselves but business-oriented informational products based on specific calculations performed on personal data. Their products are lists of qualified leads for direct marketing, marketing segments describing lifestyles and consumption patterns in general terms (e.g., Personix segmentation by Acxiom), enrichment of existing databases (adding variables to one's customers' CRM bases), and risk calculation (scoring) on individuals or profiles, particularly in the banking and insurance sectors.

Cookies and the Emergence of Data-Driven Advertising

The online advertising market is the other area in which personal data are injected into the value chain and enrich the products exchanged by market players. In the traditional media/advertising economy, the cornerstone of business relations is the audience, which acts as a currency traded between

the sector's various players: advertisers purchase access to audiences from the media, and audience measurement institutes (such as Nielsen and Comscore) measure their characteristics and thus their value (Napoli 2003).

A large part of the evolution of the online advertising sector, since its very inception, has been based on taking into account web users' browsing behavior to an ever-increasing degree, so as to qualify and estimate the value of advertising spaces, create products that match consumers' purchase intentions as closely as possible, and make money out of this profile information by selling it to other players in the value chain (Beuscart and Mellet 2013).

Today, the market infrastructure of online advertising—especially its “display” segment, which gathers various forms of advertising which appear on publishers' websites, next to editorial content and account today for about one third of online advertising spending—heavily relies on cookies. Cookies enable a website to store information on the browsers that communicate with it. They are very useful, as they mean a website does not have to rediscover an unknown web user on every page he or she visits. Instead, Internet users can have browsing sessions, store items in shopping baskets, and save their preferences. What we have seen is a redefinition of the audience product based on the tracking capabilities provided by browsing cookies. This evolution was done in three steps.

First, in the early days of online advertising, pricing and sales were based on the model used in the traditional media, with audience measurement used to qualify the profiles of visitors to a site and advertising agencies adding to those qualifications with the content of the pages viewed. The first change cookies made to this model was the possibility for a website and its advertising team to separate visits to sports-themed pages from sales of sports-themed banner ads: by tracking its visitors' browsing over time, a website can now target sports lovers on any of its pages.

As a second step, this first separation, that of visitor qualification from content viewed, was quickly joined by a second on a larger scale: ad networks such as Weborama or Criteo aggregate knowledge of web users' movements on all partner sites—those which allow Weborama to place and then enrich cookies about their visitors. As a result, an advertiser looking for sports lovers can find them on a Weborama-affiliated site whose content has nothing to do with sports: all that is necessary is for the web user to have visited another partner site linked to the sports theme.

The third major change in this market has been the development of programmatic advertising. This is based on the sale of advertising spaces through marketplaces (or ad exchanges), which use auction mechanisms that most often operate virtually in real time, hence the name real-time bidding (RTB). This coordination mechanism automates every aspect of processing: buyers' and sellers' strategies are implemented by algorithms that assess the appropriateness of displaying a banner each time a webpage is opened and seek to optimize the value of such ads. The central element is that these marketplaces are intrinsically designed to put buyers in competition with one another based on the information they have, which once again takes the form of cookies. For all market players, and especially end-buyers (advertisers and media agencies), holding or direct access to cookie bases is a strategic issue in this environment. But not all databases of cookies have the same value: the latter depends on their size (what proportion of the population is covered?), their depth (what is the precision of the qualification of the audience segments?), and their freshness. It is the combination of these three elements that determines the ability of buyers on the one hand to recognize an Internet user when an advertising impression is auctioned and on the other hand to determine the "right" bid.

Thus, in the online advertising market, the browsing data of Internet users recorded in the form of cookies have gradually become more and more important in the daily functioning of the market. For advertisers, cookie databases are accessible to a group of players who have succeeded in taking positions across the web, typically large publishers, ad networks and a few data brokers, usually in the form of buying advertising impressions targeted at a particular audience segment—for example, car purchase intention or female teenagers. The importance of individual data is such that it is reflected in the structure of prices offered by buyers on these marketplaces: at the bottom of the scale, buyers without a cookie who bid for a price close to zero; in the middle, buyers with a third-party cookie recording sociodemographic information or, better, purchase intentions; at the top of the scale, buyers with first-party information—that is, user information that was built in-house. This latter type of information serves as input to retargeting, a display advertising technique used to display advertising to people who have previously visited the advertiser's website; the ad typically displays a picture of a product the visitor was viewing earlier but did not

purchase. Retargeting, popularized by start-ups such as Criteo, is the first step in the movement of datasetization that takes first-party data as a base. That is what we are going to see now.

Datasets for Advertisers: The Invention of First-Party Data

Personal data are used in marketing in two historically distinct ways: in the form of contact details and stable and (by and large) verified sociodemographic characteristics in the world of direct marketing; and in the form of cookies aggregating browsing traces and (by and large) predicted sociodemographic characteristics in the advertising world. These two worlds now seem set to converge. But the main instigators of this evolution are not market intermediaries such as data brokers or web services companies; they are the clients of the marketing industry: advertisers.

Data Management Platforms: A Key Infrastructure for the Assetization of Customer Data

The marketing industry has been deeply transformed by digitization; private companies have gone through a similar movement since the 1980s. Private companies have not only introduced computers and databases at any level of their activity, they also are in direct contact with their customers through digital interfaces (billing supports, websites, mobile applications, etc.). Hence, like data brokers or ad networks, they are in a position to collect data. The systematization of tracking, accompanied by the standardization of customer databases, loyalty programs, and online retail, has encouraged marketing teams inside companies to consider personal data as an asset. In the 1990s, a digitized consumer database could be eventually turned to a revenue stream if it was sold to data brokers like Acxiom; in the 2010s, companies have come to consider such information as highly valuable to the point that it should not only be traded but also used as a lever for business action. From companies' point of view, this first-party data, as opposed to third-party data (audience or individual level data available to anyone ready to pay for specific audience segments), is actionable information about their customers they have accumulated and now own in various databases (purchase history, browsing data, response to marketing campaigns, location data, etc.).

Somehow, the designation and consideration of heterogeneous datasets located in various places within organizations as first-party data is characteris-

tic of the implementation of a capitalistic reasoning applied to data. Recently, the main instrument of this evolution has been the introduction and systematizing of data management platforms (DMPs) for advertisers.

In simple terms, a DMP is a database management software that can be used to collect, sort, and analyze data from a variety of sources. The supply of DMPs for marketing purposes emerged in 2011 and has developed rapidly since. The deployment of DMPs is part of a move toward what marketing professionals call “relationship marketing.” According to DMP providers, this technology should help companies develop lasting and personalized relationships with their various audiences according to their closeness or distance: customers and prospects (the foundations of customer relationship management) and suspected purchasers or unknown Internet users (the foundations of media advertising). Before DMPs, customer information was already subject to various enrichments with external data (such as segments proposed by data brokers) and calculations (segmentation, scoring) that led to specific actions toward customers (tailored mailing/emailing, special offers), but these actions remained in the field of direct marketing and CRM. DMPs are expected to go beyond this limit, bridging the gap between advertising and direct marketing, by making first-party data actionable in all kinds of direct and indirect actions with customers and noncustomers. Customer data may, for example, be useful in identifying clients and either avoiding their inclusion in a costly media advertising campaign to recruit new clients or, on the contrary, in sending those clients personalized advertising because they are at risk of defecting. The economic consequences are potentially important for companies with a portfolio of several million customers. The DMP is sold as the main cornerstone to drive business actions toward consumers in the wild, and thus as the infrastructure that turns first-party data into assets.

Leveraging First-Party Data: IDs and Algorithms

Yet, leveraging first-party data is not trivial and involves several operations that can be divided in two families: matching and calculation. The DMPs are often described by professionals as “data lakes” merging heterogeneous data sources into one single place. This aggregation process is based on the matching of identifiers among data sources, and in the DMPs, the common identifier is the cookie. Customer databases typically use the name, email address, or phone number as an identifier while the cookie serves as an

identifier to trace the browsing of Internet users. This means it is necessary to “cookify” customer data in order to onboard it in the DMP. To do so, companies can encourage customers to log in online, so that it is possible to match an email address with a web cookie. For example, the French retailer Carrefour is forcing its customers to activate their loyalty card on their website, thus setting a cookie on their computer containing their loyalty card information, which will be reused for any further visit. Another method is to use a third-party actor specializing in onboarding, who is responsible for finding bridges between lists of addresses or phone numbers, and cookies. These specialists are developing partnerships with websites where users complete forms, such as insurance products comparators. Companies such as LiveRamp (acquired by the data broker Acxiom in 2014) are specialized in this niche. The onboarding operations of customer data are only one type of basic matching operations among others performed within or besides the DMP. Once transformed into cookies, the customer data must then be matched to the databases of the advertising partners from which the advertiser chooses to launch these campaigns. The multiplication of these operations of onboarding, matching, and synchronization also results in losses and in the end, the cookie lakes feeding the DMP often cover a limited part of the advertisers’ effective clientele.

The other way to leverage first-party data is to perform computational operations on the data. These operations rely on algorithmic apparatuses. The objective here is to identify specific audience segments within the DMP and in connection with external web user information from third-party services, based on statistical proximity. A very common operation for advertisers today is to locate a segment of highly profitable customers in their client base, run algorithms to analyze their characteristics, and then launch advertising campaigns to nonclient audience segments with the same characteristics. These so-called lookalike segments are created by ad networks or big players such as Facebook and sold in the form of advertising campaigns.

Based on these two techniques, matching and calculation, first-party data are made expandable and can combine with all kinds of third-party data. The DMP is the base of these processes within organizations. It is literally, and concretely, both a data infrastructure (it is a database, or a second-level database: a database made of databases) and a valuation device (it is a knowledge device but also an action tool and a measurement tool that very precisely assesses the efficiency of different marketing actions). It

is not purely anecdotal that the metaphor of the bank, used by the World Economic Forum to describe the mechanisms of personal data storage and marketing by individuals themselves, is reemerging, this time about DMP. In the words of eMarketer, a reference market research company in the field of digital marketing: “Data is the currency of today’s digital advertising ecosystem, and everything from media planning to analytics—and every campaign execution in between—runs on it. If data is the currency of digital marketing, then the DMP is the bank” (Fisher 2013). The metaphor is not particularly detailed in the report, but the use of the image of the bank refers to that of the institution of capitalization par excellence.

Conclusion: Assets as Market Infrastructure

Today, personal data are widely understood as the counterpart of free online services in the mainstream media, with the usual claim “if it’s free, you are the product,” and in the academic field as well (Carrascal et al. 2013). But with a closer look at how personal data are forged with tracking and collection technologies, how they circulate within the advertising industry, and how they become part of products within this industry, it appears that personal data are not marketized as the product in this market, although they came to constitute a key element in its architecture. Our inquiry into the integration of personal data within the industry of marketing allows us to highlight three aspects of assetization that are more general in scope than the case described in this chapter.

First, assetization, as a process, is a combination of capture and repurposing. Interestingly, public debates on already assetized things typically focus on how firms capture things through a combination of legal and socio-technical tools to ensure rentiership from these assets (Birch 2019). For instance, Ebeling (2016) shows that private healthcare companies have elaborated sophisticated arguments to claim that patient data need to be considered as their own property. Still, this protective move comes second in the process of assetization. Before that, data-things must be considered as valuable, and this implies that they are necessarily repurposed—transposed from a specific scene to a different scene where they can be valued and eventually traded. On this basis, as revenue streams become real, their technical and legal frame is consolidated, and supplementary revenue streams are ensured through specific infrastructures that can almost cannibalize the

original “things.” Cookies are a good example of the infrastructural dimension of assetization: a simple mechanism initially aimed at filling online baskets has little by little been invested and overloaded by marketing professionals, who have built upon it complex and intricate market data-based infrastructures for data-related goods (ad exchanges, real-time bidding, etc.). Eventually, the cookie has been partly cannibalized by the ad tech perspective, which turned it into its key infrastructure for the capitalization of personal data (Mellet and Beauvisage 2019).

Second, assetization can be considered as an entrepreneurial work, especially in the repurposing stage of the assetization process. Although unsuccessful, the attempts to create a C2B market for personal data are exemplary of the dynamics of start-up entrepreneurship in the digital economy. Another example is the way the first ad networks convinced numerous websites to put tracking pixels on their pages in exchange for audience statistics, or how today firms invest in data management platforms software as a potential revenue stream for the future. In all cases, assetization is a dynamic process of exploration driven by entrepreneurs as market intermediaries and makers, with uncertain and unstabilized future revenue streams.

Finally, once stabilized, the resulting status of assets is plural and versatile. Within the online advertising market, personal data are considered as informational goods when data brokers complement consumer bases with their own information sources; they are market-based matching devices when advertisers chase their own customers on Facebook to address them with tailored messages; they are essential qualities for the valuation of ad inventories for online publishers and ad networks. In all matters, they have become essential elements in the digital advertising industry, and the common claim that digital advertising is selling user data is, although not correct in most cases, a good synecdoche for what digital ad inventories have become. However, the assetization process has led firms to consider personal data as a valuable asset for their own purpose, coining them as “first-party data.” These untraded assets have more to do with global efficiency and work instruments for marketing operations. This plurality of economic statuses demonstrates the versatile and plural nature of assets.

Note

1. ITL. Petit historique du marché de la location d'adresses. Corporate document, <http://www.itl.fr/en/fichier/59-petit-historique-marche-location-adresses>.

References

Araujo, L., and Kjellberg, H. 2016. Enacting Novel Agencements: The Case of Frequent Flyer Schemes in the US Airline Industry (1981–1991). *Consumption Markets & Culture* 19 (1): 92–110.

Beresford, A. R., Kübler, D., and Preibusch, S. 2012. Unwillingness to Pay for Privacy: A Field Experiment. *Economics Letters* 117 (1): 25–27.

Beuscart, J.-S., and Mellet, K. 2013. Competing Quality Conventions in the French Online Display Advertising Market. *Journal of Cultural Economy* 6 (4): 402–418.

Birch, K. 2019. Technoscience Rent: Toward a Theory of Rentiership for Technoscientific Capitalism. *Science, Technology, and Human Values*, doi:10.1177/0162243919829567.

Bouk, D. 2017. The History and Political Economy of Personal Data over the Last Two Centuries in Three Acts. *Osiris* 32 (1): 85–106.

Carrascal, J. P., Riederer, C., Erramilli, V., Cherubini, M., and de Oliveira, R. 2013. Your Browsing Behavior for a Big Mac: Economics of Personal Information Online. In *Proceedings of the 22nd International Conference on World Wide Web*, 189–200. ACM.

Christl, W., and Spiekermann, S. 2016. *Networks of Control. A Report on Corporate Surveillance, Digital Tracking, Big Data & Privacy*. Vienna: Facultas.

Coll, S. 2013. Consumption as Biopower: Governing Bodies with Loyalty Cards. *Journal of Consumer Culture* 13 (3): 201–220.

Danezis, G., Lewis, S. and Anderson, R. J. 2005. How Much Is Location Privacy Worth? In *WEIS* 5.

Ebeling, M. 2016. *Healthcare and Big Data*. New York: Palgrave Macmillan.

Eubanks, V. 2018. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. New York: St. Martin's Press.

Federal Trade Commission. 2014. *Data Brokers: A Call for Transparency and Accountability*. Washington, DC: Federal Trade Commission.

Fisher, L. 2013. Data Management Platforms: Using Big Data to Power Marketing Performance. EMarketer Report.

Fourcade, M., and Healy, K. 2017. Seeing like a Market. *Socio-Economic Review* 15 (1): 9–29.

- Goss, J. 1995. "We Know Who You Are and We Know Where You Live": The Instrumental Rationality of Geodemographic Systems. *Economic Geography* 71 (2): 171–198.
- Hoofnagle, C. J. 2003. Big Brother's Little Helpers: How ChoicePoint and Other Commercial Data Brokers Collect and Package Your Data for Law Enforcement. *North Carolina Journal of International Law and Commercial Regulation* 29: 595–637.
- Huberman, B. A., Adar, E., and Fine, L. R. 2005. Valuating Privacy. *IEEE Security Privacy* 3 (5): 22–25.
- Kitchin, R. 2014. *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*. London: Sage.
- McFall, L. 2014. *Devising Consumption: Cultural Economies of Insurance, Credit and Spending*. London: Routledge.
- Mellet, K., and Beauvisage, T. 2019. Cookie Monsters: Anatomy of a Digital Market Infrastructure. *Consumption Markets and Culture*, 1–20.
- Milyaeva, S., and Neyland, D. 2016. Market Innovation as Framing, Productive Friction and Bricolage: An Exploration of the Personal Data Market. *Journal of Cultural Economy* 9 (3): 229–244.
- Muniesa, F., Doganova, L., Ortiz, H., Pina-Stranger, A., Paterson, F., Bourgoin, A., Ehrenstein, V., Juven, P.-A., Pontille, D., Saraç-Lesavre, B., and Yon, G. 2017. *Capitalization: A Cultural Guide*. Paris: Presses des Mines.
- Napoli, P. 2003. *Audience Economics: Media Institutions and the Audience Marketplace*. New York: Columbia University Press.
- Nissenbaum, H. 2017. Deregulating Collection: Must Privacy Give Way to Use Regulation? SSRN, <https://ssrn.com/abstract=3092282>.
- Olejnik, L., Tran, M.-D., and Castelluccia, C. 2014. Selling Off Privacy at Auction. In *Proceedings 2014 Network and Distributed System Security Symposium*. San Diego: Internet Society.
- Posner, R. 1981. The Economics of Privacy. *The American Economic Review* 71 (2): 405–409.
- Pridmore, J. 2008. Loyal Subjects? Consumer Surveillance in the Personal Information Economy. PhD Thesis. Kingston, ON: Queen's University.
- Schneier, B. 2015. *Data and Goliath: The Hidden Battles to Collect Your Data and Control Your World*. New York: Norton.
- Schwab, K., Marcus, A., Oyola, J., Hoffman, W., and Luzi, M. 2011. *Personal Data: The Emergence of a New Asset Class*. Switzerland: World Economic Forum.
- Staiano, J., Oliver, N., Lepri, B., de Oliveira, R., Caraviello, M., and Sebe, N. 2014. Money Walks: A Human-Centric Study on the Economics of Personal Mobile Data.

In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing—UbiComp '14 Adjunct*, 583–594. Seattle: ACM.

Steel, E., Locke, C., Cadman, E., and Freese, B. 2013. How Much Is Your Personal Data Worth? *Financial Times*, June 12, 2013.

Turow, J. 2012. *The Daily You: How the New Advertising Industry Is Defining Your Identity and Your Worth*. New Haven, CT: Yale University Press.

Watson, N. 2013. *A Brief History of Experian. Our Story*. Dublin: Experian.

