Consumption Markets & Culture

ISSN: 1025-3866 (Print) 1477-223X (Online) Journal homepage: https://www.tandfonline.com/loi/gcmc20

Cookie monsters. Anatomy of a digital market infrastructure

Kevin Mellet & Thomas Beauvisage

To cite this article: Kevin Mellet & Thomas Beauvisage (2020) Cookie monsters. Anatomy of a digital market infrastructure, Consumption Markets & Culture, 23:2, 110-129, DOI: 10.1080/10253866.2019.1661246

To link to this article: https://doi.org/10.1080/10253866.2019.1661246

CONSUMPTION MARKETS & CULTURE
2020, VOL. 23, NO. 2, 110–129
https://doi.org/10.1080/10253866.2019.1661246

Routledge
Taylor & Francis Group

Cookie monsters. Anatomy of a digital market infrastructure

Kevin Mellet and Thomas Beauvisage

Orange Labs, SENSE, Châtillon, France

ABSTRACT

The aim of this article is to examine the market arrangements built by the online marketing industry around small pieces of data now ubiquitous in digital markets–namely "http cookies." We show how cookies have become the backbone and the main vehicle of a vast market infrastructure, based on its ability to transform online behavioral information into data assets, and to attach these assets to advertising products. We examine the complex trading operations that are implemented from the elementary brick that constitutes the cookie. We also raise the question of the strength and durability of this infrastructure, at a time when it is disputed and seems weakened. Beyond the particular case of cookies, we identify three main operations that market infrastructures typically support: knowledge production, capitalization, and coordination. We also highlight the centrality of "datafication" (tracking, "data lake" building, matching, etc.) in the process of market digitalization. We thus contribute to the framing of the concept of (digital) market infrastructure.

ARTICLE HISTORY
Received 15 January 2018
Accepted 14 August 2019

KEYWORDS
Market studies; digitalization; cookies; online advertising; market infrastructure; personal data

1. Introduction

> We are talking about an invention of the 1990s that the world of advertising has appropriated and which it has diverted from its original purpose. A Frankenstein monster, not really effective, but on which we have built so much value that it has always been difficult to reform it. An advertising executive, about the cookie[1]

Http cookies, or simply cookies, are small pieces of information placed and stored on web browsers by websites as users browse pages. Since their invention in 1994 by a Netscape developer, they have had a remarkable fortune: their use is almost universal in areas such as web session management, audience measurement or content customization (Schwartz 2001; Greenstein 2015). But among all the application domains of cookies, online advertising is certainly the industry that has designed the largest and the most sophisticated cookie-dependent architecture (Turow 2012; Smith 2014). Over the past twenty years, this industry has come to rely more and more on cookies, to the point of becoming totally addicted to them[2]: the commercialization of advertising inventory (especially in the sub-field of display advertising[3]), and the coordination between the multiple intermediaries that make up the online advertising food chain, is very much based on the mobilization

CONTACT Kevin Mellet kevin.mellet@orange.com Orange Labs, Sense, Orange Gardens (2B), 44 avenue de la République, Châtillon 92320, France

[1]Vincent Tessier, Vice President of ad-tech firm Adsquare, in Jaimes N., 2018, « Comment l'adtech peut-il survivre au monde post-cookie ? » [tr: how can adtech survive in a post-cookie world?], Journal du Net, September.

[2]Hence the reference to Cookie Monster in the title of this article. Cookie Monster is a muppet from the TV show Sesame Street, well known for his taste for cookies. The gluttony of Cookie Monster, and his particular taste for cookies, make him the perfect symbol of the online advertising industry, which as a whole has been characterized in recent years by its fierce appetite for online tracking, and by its immoderate taste for cookies. This reference to popular culture has not remained foreign to the world of technology and advertising, since Cookie Monster is the name given to a web extension that automatically erases cookies; it is also the name of the core component of major advertising technology firm AppNexus' infrastructure, its huge server-side cookie database.

and circulation of cookie-based information (Englehardt and Narayanan 2016). In addition, some so called "third-party" players have managed to form relationships with large networks of websites, allowing them to build databases recording information on virtually any Internet user (Falahrastegar et al. 2014; Binns et al. 2018). It is common that the mere opening of a page on a website generates the production or updating of several dozen cookies. And advertising market players themselves come to observe that cookies proliferate, and that this proliferation might be a problem in itself (IAB 2014).

This article describes the process of "cookification" of the online advertising industry. Following a science and technology studies (STS) perspective, we show how this innocuous piece of code has become the backbone and the main vehicle of a vast market infrastructure, based on its ability to transform online behavioral information into data assets, and to attach these assets to advertising products. We examine the complex trading operations that are implemented from the elementary brick that constitutes the cookie. Finally, we raise the question of the strength and durability of this infrastructure, at a time when it is disputed and seems weakened by the growth of mobile uses and the establishment of new regulations. Beyond the particular case of cookies and the advertising industry, we show the interest to pay particular attention to market infrastructures, basically defined as the material and discursive objects and assemblages that create the grounds on which markets silently operate. Following recent conceptualizations by Chakrabarti et al. (2016) and Kjellberg, Hagberg, and Cochoy (2019), this article seeks to answer the following two questions: What does the particular case of cookies bring to the general understanding of market infrastructures? What lessons can we draw from the point of view of the study of the digitalization of markets?

Our argument is that the history of the "cookification" of the online advertising industry illustrates key properties of digital market infrastructures. In many ways, "cookification" appears to be a very traditional process of infrastructuring, the success of which is measured by its ability to enroll new players, to absorb old devices and mechanisms, and to foster new uses (Star and Bowker 2002). Today, there is no business, online or brick and mortar, which escapes the influence of cookies. The ubiquity, invisibility (in the eyes of ordinary Internet users, at least), and permanence of cookies contribute to its status as a market infrastructure (Kjellberg, Hagberg, and Cochoy 2019).

Our study allows us to go further into the elaboration of this conceptual framework. In the fields of marketing and advertising, cookie databases have a special function: they allow consumers to be described from the recording, aggregation and analysis of their web browsing activity, within a web site or from site to site. Cookies thus make it possible to produce a limited but extensive knowledge of Internet users. This knowledge is stable and transferable, through time and space, because the cookie contains a unique identifier that allows whoever manages it to find a user again on another website.[4] Through the multiple matching and aggregation operations that it allows, for example when cookies are pooled into audience segments, this identifier serves as a powerful means for operations of capitalization and valuation by the various market players involved in the online advertising food chain. Finally, cookies allow marketers and advertizers to "follow" users and to address them with various forms of targeted marketing and advertising. In other words, the cookie is at the center of a sociotechnical system that feeds a knowledge base produced from the tracking of consumers, the capitalization of this knowledge into marketable segments, and its operationalization in advertising and marketing actions. We argue that the combination of these operations of knowledge production, capitalization and coordination, turn the cookie into a very powerful market infrastructure.

The article is structured as follows. Section One presents our theoretical framework and research objectives. We clarify our use of the concept of market infrastructure, which occupies a central place in our approach. Section Two outlines our data sources and methodology. The next two sections describe the "cookification" of the online advertising industry, through the history, the effects and

[3]Display refers to several forms of advertising which appear on publishers' websites, next to editorial content: banners, text ads, rich media, video ads, etc. (see Beuscart and Mellet 2013).

[4]This identifier gives the cookie the status of "personal data" within privacy laws.

the legacy of the cookie infrastructure. Section Three describes the formation of a cookie-based advertising market infrastructure by telling the different steps that led to the widespread adoption of cookies by advertising professionals. We analyse how the advertising industry has opportunistically succeeded in integrating the cookie in a very diverse set of activities, and thus contributed to turn it into a market infrastructure. Section Four raises the question of the solidity and sustainability of the cookified market infrastructure as it is challenged. In doing so, we seek to characterize and examine the causes of the power of cookies as a market infrastructure. Finally, we provide a summary and conclusions.

2. Theoretical framework and research objectives

To what extent can an analysis of cookies, and the ongoing process of "cookification" of online advertising, shed light on the shaping and operation of digital markets? In this section, we present our theoretical framework, focused around the concept of market infrastructure. First, we provide a brief survey of the literature on the movement of "datafication," with a focus on works that are concerned with marketing and advertising. Then we are interested in the concept of market infrastructure. We recall the genealogy of the concept and specify how we contribute to its theoretical development.

2.1. Datafication and the marketing industry

Big data technologies, and especially those using personal data, have received growing academic attention in recent years. Many studies have investigated the social and economic consequences of digital knowledge infrastructures, tracking devices, and algorithms, using the term of "datafication." It is possible to identify two main perspectives inside this corpus.

A first set of works adopts a wide-angle lens and observes a largely coherent, intentional and convergent dynamic, towards the setting up of an infrastructure mainly dedicated to tracking and surveillance. Schneier (2015) points out that, "in the normal course of their operations, computers continuously document what they're doing" (15); big data systematizes the apprehension of data and traces as a valuable resource requiring attention and investment. Kitchin (2014) describes how the manufacture of data is greatly linked to their use:

> databases are designed and built to hold certain kinds of data and enable certain kinds of analysis, and how they are structured has profound consequences as to what queries and analysis can be performed; their formulation conditions the work that can be done through them. (…) Data infrastructures host and link databases into a more complex sociotechnical structure. (21–23)

This sociotechnical structure is the subject of a vast body of research that shares a common description of datafication as a convergent process of constitution of a vast tracking and surveillance infrastructure (Lyon 2001; Zimmer 2008; Christl and Spiekermann 2016; Zuboff 2019). These works take a critical look at this movement and denounce the risks they pose to individual freedom and privacy. This dynamic, and the controversies and socio-economic struggles in opens, are of course central. But the focus chosen by these authors tends to leave aside the specific application fields and concrete uses of these tracking systems.

A second set of works on datafication is more focused on the multiples uses of data by businesses in the practical course of their activity, especially in the area of digital marketing. These studies elaborate a more disordered, and largely unintended narrative, which leads to the recognition of the omnipresence of data in the social and economic existence of the digital economy.

As markets are digitalized, data are a key element in the operation of markets in a variety of forms: consumer records, scoring and targeting instruments, customization engines, algorithmic pricing, vending and trading machines. Hence, data act as a versatile, yet vital, apparatus for markets, through the construction of data as a device, a valuation tool, a measure, an asset, or an

infrastructure. Following this perspective, Cluley (2018) defines the process of datafication as the non-marketing activities needed to create a standard unit of data. From a case study on the implementation of a measure of the viewability of advertisements on the Internet, he describes a three steps process, "where an alignment of interests stabilizes an object of measurement which is then formalized in a standard method for measurement."

Although data collection practices for marketing purposes are hard to investigate (see below our methodological framework), attempts to assess them, or to report on the data brokerage industry, have come to constitute a cumulative knowledge base on the sociotechnical market arrangements built upon personal data. Based on a self-ethnography, Ebeling's work (2016) shows how ordinary life events become objects of capitalization by data brokers, private companies that manufacture and maintain databases of "digital doubles" of virtually each individual consumer. These data doubles exist without the knowledge of and away from their human alter egos, and come to haunt them, more and more frequently, like phantoms (or monsters!), in the form of hyper-targeted marketing. Data doubles are enrolled in a set of economic activities, and valued as commodities or assets (Bouk 2017; Neyland and Milyaeva 2016; Fourcade and Healy 2017), or exploited as segments and marketable groups (Turow 2012; Ariztia 2015; Cluley and Brown 2015).

This disparate body of research shows how data has become a central concern for people and organizations operating and acting on markets. Yet, these works have only paid oblique attention to the materiality of these data doubles, especially to their cookie-based nature–with the notable exception of Turow (2012) and Cluley and Brown (2015). Turow's work is focused on the advertising food chain and its successive evolutions guided towards more personalization and targeting. Cluley and Brown investigate the development of cookie-based market devices, and focus on how such technologies change the way consumers are represented, categorized and, finally, "dividualized" for marketing operations as fine-grained ad-hoc segments and specific properties.

Finally, the http cookie is also an object of interest in computer science, with works studying the mechanisms of production and management of cookies, and technical coordination between actors having use of cookies (Park and Sandhu 2000; Mayer and Mitchell 2012; Ghosh et al. 2015). Cookies have also been of interest to lawyers and privacy scholars because their legal status as personal information appears fluctuating and uncertain (Siebecker 2002; Mills 2008). In addition to this scientific literature, there are more personalized accounts of the invention and development of cookies, which place particular emphasis on the hesitations, decisions and remorse of their inventor (see in particular Schwartz 2001).

2.2. From market devices to market infrastructures

The study of market datafication can benefit from two neighboring fields which share a common affiliation with STS, but which have undergone separate developments: infrastructure studies, and "market studies." Infrastructure studies were especially theorized as a research agenda by Star (Star and Ruhleder 1996; Star 1999), and Bowker et al. (2009). Following a socio-historical perspective, these authors set themselves the objective of putting into the foreground what, in general, is in the background, invisible and transparent. Infrastructures are described as a vast set of objects that are necessary for most human activities: material collective equipment, such as bridges, energy or communication networks, but also inscriptions such as protocols, standards and classifications. There is now a vast literature in infrastructure studies, from which there are several interesting observations for our purpose. Good and usable infrastructures tend to disappear, "by definition" (Bowker and Star 1999), which makes them particularly hard to see. Infrastructures are often in the making and expanding, and they rarely result in their actual form of a defined and centralized plan. Standards and classification are of crucial importance and they embed social and political arrangements and compromises.

The other approach has developed from the work of Callon (1998) in economic sociology and then expanded to related disciplines such as marketing and consumer research (Geiger, Kjellberg,

and Spencer 2012; Cochoy, Trompette, and Araujo 2016). It has contributed in a number of ways to the renewal of market understanding, highlighting the role of objects in economic coordination. Of particular interest for our analysis is the concept of "market devices," defined as the material and discursive assemblages that intervenes in the construction of markets (Muniesa, Millo, and Muniesa 2007). This opens up an analysis of the diversity of market devices and, beyond, of market agencements (Callon 2016), whether apprehended by the way devices frame the cognition and calculability of participants, or by the way they contribute to the valuation of goods and to the formation of prices.

These two perspectives are not at all incompatible, but they tend to focus on different things: on one side, large ensembles that raise problems of maintenance, control, or interoperability; on the other, material things that frame market coordination (trolley, formula, matching algorithm, website interface, etc.). On the one hand, invisible infrastructures, technical systems, classifications and standards that create the ground on which social life takes place; on the other hand, material and discursive devices that are visible and play a role of intermediary in the economic coordination.

2.3. Research objectives

In this paper, we wish to contribute to the sociology of markets by framing the concept of *market infrastructure*, based on the case of the cookie in the online advertising industry. At the intersection of infrastructure studies and market studies, the concept of market infrastructure is a good candidate to describe the set of objects that create the grounds on which markets operate–to paraphrase Larkin (2013, 329). Recent theorizations by Chakrabarti et al. (2016) and Kjellberg, Hagberg, and Cochoy (2019) already combine market studies and infrastructure studies, and constitute a precious aid for our investigation. Chakrabati et al. introduce the concept of market infrastructure, and look at how market infrastructures contribute to shape and stabilize markets. Kjellberg et al. define the market infrastructure as "a materially heterogeneous arrangement that silently supports and structures the consummation of market exchanges." Based on a history of barcodes in retail, they propose to identify a non-exhaustive list of eight properties associated with market infrastructures: market infrastructures are relational, available for use, modular, actively maintained, interdependent, commercial, emergent and political. These authors also underline the interest of the concept of market infrastructure as a way to describe the arrangement and combination of market devices, and to render the specific agency and political dimension of invisible market infrastructures.

We would like to contribute to this nascent reflexion on market infrastructures by stressing the key market operations they support. Infrastructures are background resources that simultaneously sustain and frame the actors' actions. Hence, the definition of market infrastructures should not (only) focus on what they are made of, i.e. things, relations, standards, norms etc., but rather on what kind of actions or activities they sustain for the performance of market activities. We formulate the hypothesis that market infrastructures support three main operations. First and foremost, market infrastructures are knowledge infrastructures: they enable the production and circulation of information useful for economic transactions. This is the proper Hayekian function of markets (von Hayek 1937, 1945), and a dimension well put forward by the contribution of infrastructure studies (Bowker et al. 2009; Edwards et al. 2013). Second, market infrastructures are capitalization, and by extension, valuation infrastructures (Muniesa et al. 2017; Birch and Muniesa forthcoming). Infrastructures are sets of resources (more or less) accessible and available to market participants, or to certain sub-groups of market participants. Securing access to these resources can generate revenue streams, rents, and sometimes market power. This is why the more or less open nature of access to infrastructures, and all the upstream work of standardization, is of a strategic nature for firms (Brunsson and Jacobsson 2000). Third, and quite obviously, market infrastructures are coordination infrastructures: they make interactions between market participants, in the form of exchanges and transactions, possible, and *in fine* they make their individual decisions compatible (Arrow 1974; Simon 1991). These infrastructures are material but they also require a set of conventions to reduce the uncertainty inherent in the coordination of individual actions (Favereau and Lazega 2002).

These three operations (knowledge production, capitalization, coordination) are necessary for the proper functioning of markets. They can be realized and supported by various devices, although the distinction between devices and infrastructures is not always obvious. Online consumer reviews are a good example (Mellet et al. 2014). This digital market device makes available to consumers a new form of knowledge, that complements judgment devices already available on markets (Karpik 2010). It may be the subject of capitalization strategies by firms seeking to maintain and strengthen their reputational capital, but such strategies are far from systematic (Beuscart, Mellet, and Trespeuch 2016). Yet, online consumer reviews do not play a direct role in the coordination of supply and demand and in the fulfillment of market exchanges. Thus, online consumer reviews are a good example of a market device, but not of a market infrastructure. We will therefore talk about market infrastructures when operations of knowledge production, capitalization and coordination are carried out, and encapsulated in socio-technical systems that function silently and ubiquitously in the background of the markets. This triple condition is sufficiently demanding that any market device cannot be assimilated to a market infrastructure. It is though perfectly suited to the barcode, studied by Kjellberg et al., and to cookies, as we will try to show.

3. Methodological framework

Studying infrastructures raises an array of challenges, as already pointed by Star (1999). As support objects for the activity of various kinds of actors, many complementary approaches can be used to describe infrastructures and their effects: observing social activity in committees and meetings where standards and technical specifications are decided, analysing infrastructures when they are in crisis or collapse, accounting for their materiality (road length, production capacity, etc.), unveiling their systemic dimension by studying standards, classifications or relationships between actors that form networks, etc.

Recent works, relevant to our study, build upon the analysis of discussions, deliberations and decisions in the framework of committees, industrial projects or professional organizations, with the aim to define the standards and technical specifications supporting the establishment of market infrastructures (Kjellberg, Hagberg, and Cochoy 2019) or market measurement devices (Cluley 2018). This type of method is not suitable when analyzing cookies, simply because such collective bodies did not exist. Admittedly, discussions took place within the framework of the Internet Engineering Task Force (IETF), and Lou Montulli, inventor of the cookie, attended them (Schwartz 2001; Peacock 2014). But these discussions focused mainly on the impact of cookies on privacy, and the decisions they made were never followed. The cookie-based market infrastructure has emerged in a silent way, outside of any discussion, even among advertising professionals.

When studying digital infrastructures, the material dimension seems hard to document: of course, digital infrastructures are made of computers and servers, electricity, wires, etc. However, as "networks of networks" (Edwards 1998), digital infrastructures consist in intricate hardware and software, back office and user interface layers. In this regard, studying the material dimension of the online advertising industry's infrastructure is also a matter of software environment: how do web browsers work, which protocols do they implement, what is inside an ad server, etc.? The methodological difficulty here rests upon the necessity to understand technique as an array of possible and impossible things, a set of existing frames that can be used, tweaked or influenced by the ad industry. Christl and Spiekermann's (2016) detailed listing of tracking devices and methods is a valuable example of such approach. The necessity to build upon a kind of technical literacy is reinforced by the opacity of the ad industry and the practices related to manipulating personal data. In its 2014 report on the data brokerage industry, the US Federal Trade Commission underlines that this sector is complex, and that brokers "collect consumer data from numerous sources, largely without consumers' knowledge" (FTC 2014). Reasons for this opacity are multiple: it is a complex ecosystem,[5] it is moving fast, it is a highly competitive and competitive advantage is mostly built upon proprietary (that is secret) technology, fraud is common, major players do not talk, etc.

As result, our investigation on the cookie-based market infrastructure of online advertising combines multiple approaches in order to build a full understanding of it as a sociotechnical object. First, we had a close look at the technical equipment of the online ad industry: how browsers and cookies work, but also the history of the technical mechanism built upon them for advertising, i.e. ad servers, ad exchanges, data lakes, etc. Such investigation relies on the reading of technical documentation, online blogs dedicated to the field, and discussions between professionals. Second, we scrutinized both lay and academic inquiries on the data brokerage industry: investigations conducted by journalists, but also original experiments in the field of computer science that assess how the ad industry works and prices things (Castelluccia, Grumbach, and Olejnik 2013; Olejnik, Tran, and Castelluccia 2014; Englehardt and Narayanan 2016; Brookman et al. 2017; Binns et al. 2018). Third, we observed how professionals themselves report on their practices, their technical environment, and the evolution of their industry. Such investigation involves the attendance to and recording of conferences and professional meetings, and the reading of the trade press. Finally, we interviewed professionals during investigations conducted between 2010 and 2018 on various aspects of online advertising: ad exchanges, the assetization of personal data, targeting categories (see Beuscart and Mellet 2013; Beauvisage and Mellet forthcoming). Cookies were not the main topic of these interviews, but appeared spontaneously as a recurring concern for interviewees.

4. The rise of a market infrastructure: how online advertising embraced cookies

In this section, we describe how a vast market infrastructure has formed around the cookie. After a brief return on the emergence of the cookie, in the 1990s, we highlight the three steps followed by the advertising industry to adopt the cookie and use it for operations of knowledge production, capitalization, and coordination – three properties that contribute to turn cookies into a market infrastructure.

4.1. The invention of the cookie

The cookie was originally designed in 1994 by Lou Montulli, a member of the Netscape browser development team (Schwartz 2001; see also Turow 2012). The team was grappling with how to respond to a problem that online commerce sites encounter in shopping cart management: at the time, there was no satisfactory technical solution to put several items in a shopping cart before paying for purchases. The cookie lets the browser record lists of items, information on the purchase route, the content of forms, allowing sites to resume a purchase route interrupted several days before.

The mechanics of cookies is relatively simple: they allow a website to store information in the browser of their visitors. Each visited site can thus create a small local database, in the form of a data dictionary, a list of variable-value pairs, which it can create, modify and consult each time a visitor consults one of its pages. For each page visit, the browser sends the server the content of the cookie concerning it, and the server can adapt the page returned and update the content of the cookie. The cookie is a major advance for websites. It responds very well to the primary need of basket management, and generally allows building much more interactive and personalized content. Its advantages do not stop there: the cookie also makes it possible to measure the visit of a site more accurately than before. Barely two months after their creation, cookies were integrated in version 1.0 of the first commercial browser, Netscape Navigator, distributed from December 1994. Internet Explorer, Microsoft's competing browser, incorporated the cookie from its version 2 released in October 1995. Browsers were set to accept cookies by default, and cookies were designed

[5]A common representation of this complexity is the so-called "Lumascape," a diagram that circulates widely among advertising professionals to describe the programmatic value chain, the different types of intermediaries involved, and the myriad players involved.

to exchange silently, without alerting the user–the latter could however delete cookies, through the browser's settings.

How could this elementary mechanism, based on the storage of information by a site on a user, interest the advertising industry and be at the center of debates on the privacy of Internet users? This is due to the combination of the cookie mechanism with another fundamental mechanism in the design of web pages. A web page is not a single and monolithic document, a single file; it is a composition of visible and invisible elements assembled by the browser: text, images, video, discussion threads, code, style sheets, etc. (Mayer and Mitchell 2012). In each of these cases, the browser may call a server other than the one that manages the web page the user is consulting, a "third-party" server that can both, display specific elements, and create, read and modify one's own cookies. We can thus distinguish between *first-party cookies*, those managed by the web site (of the domain and its subdomains) the user is consulting, and *third-party cookies*, relating to servers called by the page that the user is consulting. This subtlety was not anticipated by the cookie's initial developers, and Netscape's late attempts to allow users to set their third-party cookies did not take hold (Schwartz 2001). Third-party cookies, soon adopted by the advertising industry, have turned the overall cookie mechanism into a general tracking system, thus perverting the initial objectives of its designers. In a late interview, Lou Montulli clearly stated that "we didn't want cookies to be used as a general tracking mechanism" (Schwartz 2001).

4.2. The early days of cookie-less online advertising

The "cookification" of the advertising industry was not immediate: it involved changes in the way advertizers and agencies buy advertising and plan their campaigns, and in the way online publishers sell and value advertising space. The transformation was therefore slow: it started from the existing model of the advertising industry as it existed for print media or television, and gradually deviated from it. In the early days of online advertising, ad serving and pricing were based on the model used in traditional media (Napoli 2003; Beuscart and Mellet 2013). A media's audience is measured through panel data by audience measurement firms like Nielsen, and the output is both a quantification of the number of readers, viewers, visitors, etc., and a qualification of their profile. This qualification is based on a small set of variables: age, gender, income, interests–so that it can fit the very large array of advertizers' perspectives, and also be measurable using panel data. Based on this qualification of its audience, a media lets its ad sales team or a representative advertising network market ad space as part of bundles of targeted segments. After discussing and negotiating with the advertizer's media agency, in order to set the objectives of the campaign in terms of target, schedule and price, the ad network takes the responsibility to operate the campaign. On the whole, the system is large-grained: segments are large, the means to target them are made at a media or a section (sports, economics, fashion, etc.) level, and no individual measurement of exposure and its effects

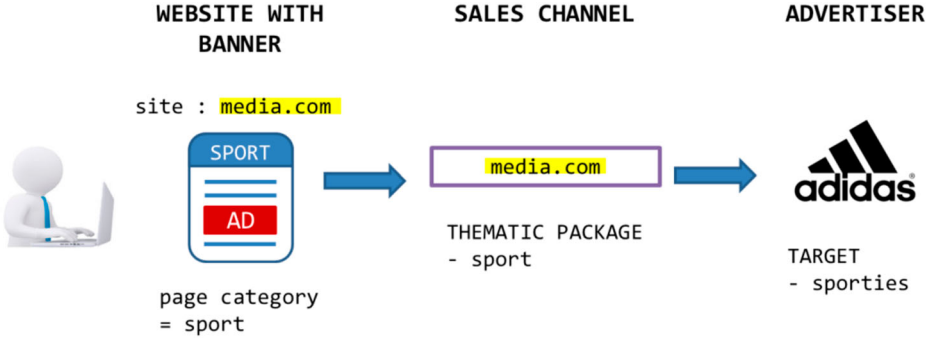

Figure 1. Deploying the traditional media-based model to online media.

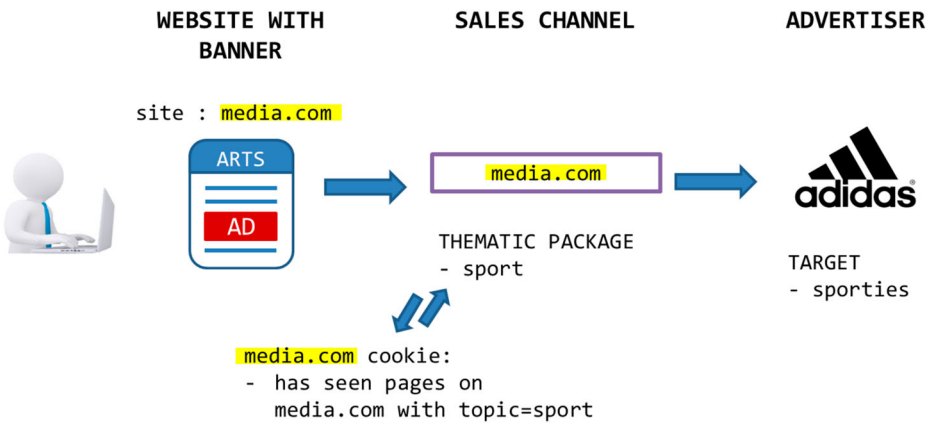

Figure 2. Online media use cookies to find targets anywhere on their websites.

exists at a global scale. The first attempts to develop online advertising inherited from this model, and tried to copy it (see Figure 1).

4.3. Step one: knowledge with first-party cookies

The first change cookies made to this model, in the late nineties, leveraged the tracking abilities provided by first-party cookies. Cookies provided media websites and ad sales the possibility to separate visits from content: by tracking its visitors using first-party cookies, a website could target sports lovers on any of its pages (see Figure 2). Cookies brought a softening of constraints for the delivery of campaigns for specific targets, and the ability to handle users' targeting at a whole domain level. With this new user-centric way of matching marketing segments and web users, the cookie led to the birth of profiling and behavioral targeting. Ad sales workers still had to fulfill campaign objectives expressed in terms of profiles, delays and target quantity, but they had new means to address audiences based on their browsing behavior. Such transformation was not trivial, since behavioral information is plural and not necessarily explicit: in a context where all the users in a household use the same computer, a single cookie can visit fashion and automotive contents; campaigns' objectives can be overlapping; etc. Hence, this first step of the cookification of the advertising industry had a lot to do with experiencing and domesticating browsing information. But the whole architecture of the market remained stable: direct sales between the agency and the ad sales team prevailed, and tracking was limited to one media, so that a niche media or a website with a small audience had limited options to benefit from the cookification of its audience.

4.4. Step two: capitalization with third-party cookies

The second step of the cookification of the advertising industry was the transformation of single-media audiences into networked audiences in the first decade of the twenty-first century. By allowing third parties to build a new, valuable knowledge on web users, the mechanism of third-party cookies extracted value encapsulated inside cookies. This evolution came from third-party players which had specialized in audience measurement (like US based company Quantcast or its French counterpart Weborama). The latter offered websites to provide them with free audience data, in exchange for permission to deposit a third party cookie on their visitors' computers. The deal was interesting and inexpensive: it allowed these actors to cover a large part of the web (deals with thousands of websites, concerning millions of users). Here the interest of third-party cookies was clear: it enabled the tracking of Internet users not only inside a given website, but from site to site, which led to the

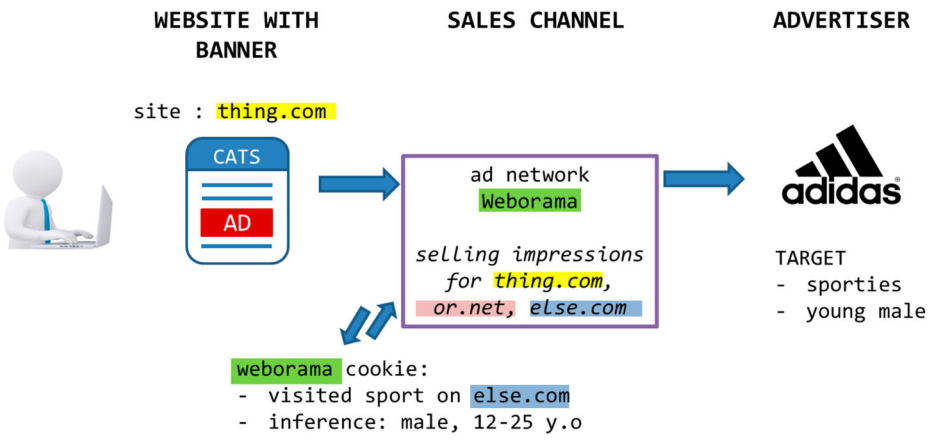

Figure 3. Ad networks capitalize knowledge on users on multiple websites, and reinforce targeting abilities.

fabrication of a statistical knowledge of these users across these sites: interests, inferred socio-demographic properties, etc. Actors such as Weborama understood that they could exploit their links with many websites and their knowledge of Internet users by transforming themselves into ad networks. They offered websites a way to sell ad space, by bundling their cookified visitors into specific audience segments. An advertizer looking for sport lovers could find them on a Weborama-affiliated site whose content has nothing to do with sport: all that is necessary is for the web user to have visited another partner site linked to the "sport" theme (Figure 3). In the end, the specialization of actors in the production and management of third-party cookie databases has contributed to transform these knowledge sets into valuable economic assets.

4.5. Step three: coordination with ad-exchanges

The third major step in the movement of cookification has been the development of programmatic advertising in the 2010s, based on the sale of ad space through marketplaces (ad exchanges), with auction mechanisms in real time (hence the name real time bidding, or RTB). This coordination mechanism fully relies on the ability of the cookie to bring together buyers and sellers of ad spaces and to set prices, based on the knowledge each buyer has accumulated thanks to the cookie technology. Ad exchanges automate the process of advertising inventory purchasing: buyers' and sellers' strategies are implemented by algorithmic mechanisms which assess the appropriateness of displaying a banner and determine the optimal price.

The central element is that these marketplaces are intrinsically designed to put buyers in competition with one another based on the information they have about the webpage's visitor, which once again takes the form of cookies. To do so, they rely on a mechanism called cookie matching, deployed by ad exchanges, whereby the cookies that each ad network has set on the user are synchronized with the ones built by the ad exchange: even without being a partner of the site currently visited by a consumer, any potential bidder can project all the knowledge it has of this consumer to buy (and possibly resell) the advertising space put up for sale on the ad exchange. The mechanics of ad exchanges extends enormously the agency of buyers in the online advertising market, and the links connecting them with Internet users. First, the various bidders may rely either on information they already have or on information acquired from data providers. Data providers such as Exelate, Bluekai, or Addthis, collect, enrich, combine and resell cookies. Like ad networks, which some of them used to be, they cover large parts of the web (Bergemann and Bonatti 2015). Second, buyers with first-party data, that is, user information that was built in-house, are able to leverage this specific information while bidding on ad exchanges. For example, browsing information on an online commerce website may serve

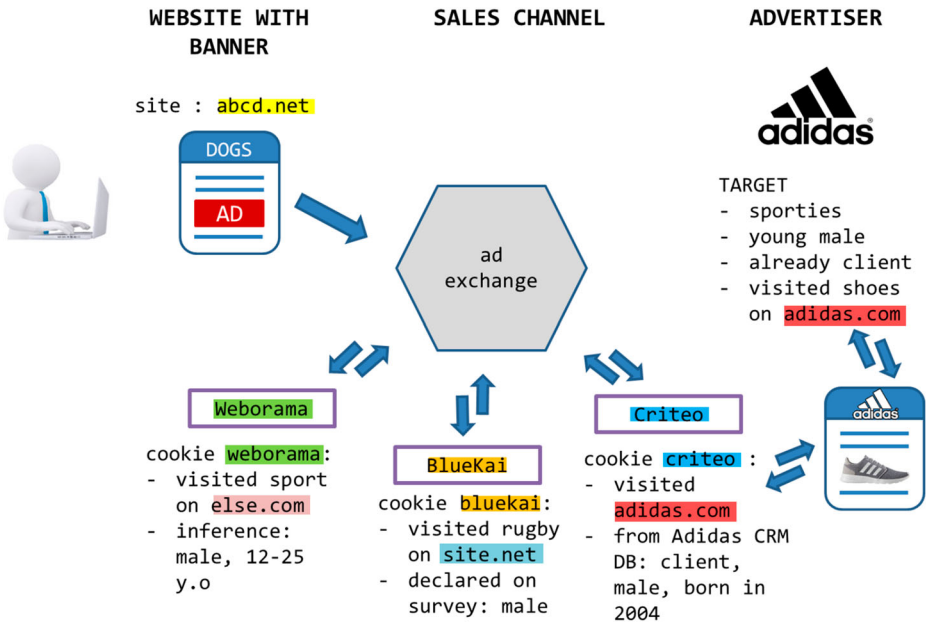

Figure 4. Ad-exchanges operate cookie matching, so that many ad networks can buy impressions from any website based on the knowledge they built on their own network.

as an input to retargeting, an advertising technique used to display advertising to people who have previously visited the advertizer's website; the ad typically displays a picture of a product the visitor was viewing earlier, but did not purchase. Retargeting was popularized by start-ups such as Criteo (Figure 4). Finally, cookies are used by advertizers on their websites to "onboard" their visitors, that is to recognize past clients among anonymous visitors: companies such as Liveramp, a subsidiary of Acxiom, a data brokerage company, are specialized in this type of service. The mechanism for synchronizing cookies, which is the essence of the work of ad exchanges, allows both individual inventory buyers and sellers to determine their bids, and the central agent to operate market clearing (Ghosh et al. 2015). Cookies are ultimately the basic brick of coordination for programmatic advertising which accounted for about 85% of digital display advertising spending in the UK in 2018 (eMarketer).

This presentation is of course very schematic. Indeed, we clearly distinguished the operations of knowledge production, capitalization and coordination, even though they are implemented simultaneously. Cookies are primarily a knowledge operator on consumers whose browsing behavior is tracked and recorded. Assembled in vast repositories, and attached to advertising products, cookies serve as support for capitalization operations. Engaged in sophisticated automated bidding and purchasing mechanisms, cookies serve as support for matching, synchronization and ultimately coordination between the multiple actors involved. As a result, the cookie has come to personify the market infrastructure of the online display advertising industry, so that today virtually no one advertisement on the web is displayed without involving the use of cookies in one way or another, and the various players of the advertising food chain are like cookie monsters, as they continuously gulp down and spit out massive amounts of consumer data in the form of cookies.

5. The power of cookies

The large range of actions that can be performed thanks to cookies fits very well the need of the advertising and marketing industry. However, surprisingly, the cookie-based market infrastructure

was not built initially for advertising purpose, and the very inventor of the cookie technology himself takes a critical look at how his creation turned into a massive tracking system (Schwartz 2001). Even actors occupying a central position in the advertising industry call it a Frankenstein monster (see epigraph). What makes the cookie so monstrous in the eyes of these witnesses? Where do the cookies derive their power from? How can one explain the digital gluttony of the advertising and marketing industry that has built a huge machinery to manufacture cookies, put them into motion and process them? What are the limits to the expansion of this huge infrastructure?

We discuss three aspects of the "power of cookies" in this section. First, we focus on the dynamic of continued expansion of this infrastructure, made up of unexpected junctions and heterogeneous attachments. Secondly, we investigate the future of the cookie while this infrastructure is questioned and weakened–in particular by the evolution of the regulation on personal data. Thirdly, we show how the cookie-based market infrastructure serves as a model in cookie-less environments.

5.1. Spreading advertising cookies all around the web

Cookies are descriptions of otherwise particularly elusive entities: consumers. Admittedly, cookie databases do not contain consumers, their changing affects and desires, their complicated lives, in their entirety, but they contain stable and transferable forms of knowledge. It is because they contain little information–an identifier and associated tags and/or scores–that cookies can travel fast and far, that they can be assembled by millions, recombined, and injected into sophisticated commercial devices such as ad exchanges.

The transformation of the online advertising industry described in the previous sections is also a change in scale: the transition from a small scale practice, reduced to websites and their servers, to the manipulation at a very large scale of behavioral data. This change in scale is difficult to measure, although some have tried to. Several quantitative analyses rely on browser extensions to detect and assess the extent of third-party tracking on websites. In a longitudinal study of tracking on a sample of 1200 popular US websites, Krishnamurthy and Wills (2009) show the increasing aggregation of user-related data (in the form of cookies) between 2005 and 2008. The top-ten third-party trackers–including those managed by Google, Yahoo, AOL or Microsoft–were used by 40% of web sites in 2005, but had extended to 70% of popular websites by 2008. In an examination of more than 144 million webpages downloaded in more than a dozen countries, performed in Dec. 2017, a team at privacy browser extension Ghostery found that 77.4% of all websites had at least one third-party tracker (MacBeth 2017). In addition, the analysis revealed that a significant share of examined websites (16.2%) featured 10 or more trackers. Google and Facebook were the most prevalent users of trackers, for analytics and advertising purposes, according to Ghostery. Thus, a handful of companies are able to track users' movements across almost all of the web sites (see also Castelluccia, Grumbach, and Olejnik 2013; Falahrastegar et al. 2014; Yu et al. 2016). Symmetrically, most dollars allocated to online advertising fuel economic circuits that use cookies in one way or another: upstream to measure the audience, downstream to evaluate the effectiveness advertising, or in the middle to identify targets and design audience segments. In other words, the scale of cookie tracking is almost as vast as the scale of the web, its web sites and all of its users.

Moreover, the use of cookies by the advertising industry does not limit to user tracking. Cookies support a large variety of operations related to daily advertising operations: capping, in order to measure and eventually limit how many times a user has been served a given ad; attribution, in order to share the revenue from a sale among the different stakeholders that have shown the ad to the user; stored ad preferences, even, ironically, to remember opt-out from targeted ads at users' request; producing analytics on pages supporting ads. The cookie has been so convenient for ad tech companies to store and exchange information on web users that it has been used as an ingredient for a large range of purposes.

The cookie infrastructure has become an essential facility for participants in the advertising market. We have seen how certain actors have patiently managed to collect traces of the behavior of

millions of Internet users, to gather them in databases, to recombine them and mobilize them in their operation. This extension of the realm of cookies to almost all the web relies on a series of enrollment and "intéressement" (Callon 1984) dynamics. The new firms specialized in online advertising (ad networks, data providers, ad verification solutions, audience measurement companies, etc.) have propagated the cookie to the two ends of the market: website publishers and advertizers. As a first step, ad networks have propagated advertising cookies to a large number of websites by offering them free audience measurement services. In the early days of the web, when audience figures and analytics were costly to acquire, ad networks offered publishers free knowledge about their audience, based on the network's web beacon put on their partners' pages. In addition to providing audience ratings, the network could capitalize on the knowledge gathered on consumers to start selling ads. Another example of such bargains is provided by AddThis, a company that provides website publishers with button solutions for sharing their content on social media sites. The success of these freely available buttons allowed AddThis to multiply the connection points with Internet users and thus to deploy a vast tracking net. Starting in the 2010s, onboarding and data management platform providers have convinced advertizers to "cookify" their transactional databases and first-party data, thus expanding the realm of cookies inside organizations.

In the well-named classic book *A sociology of monsters. Essays on power, technology and domination*, J. Law is interested on the issue of power and how the STS perspective can fuel understanding of power mechanisms. For him,

> no one, no thing, no class, no gender can 'have power' unless a set of relations is constituted and held in place: a set of relations that distinguishes between this and that (distribution) and then goes on to regulate the relations between this and that. (Law 1990, 18)

Technical artifacts, machines and inscriptions play a central role in these mechanisms as they constitute, reorganize, dissolve what sociologists traditionally call social relations. The effects of power are to be found in the mechanisms of reproduction, overlap, and stabilization (or freezing) that result from the constitution of large scale networks. This is what we observe with cookies. There is not an instance that globally controls, organizes, coordinates or modifies the properties of cookies and of the infrastructures that sustains them. On the contrary, we observe that a myriad of entities reproduce mechanisms of tracking, targeting and matching, on and on, again and again. The power of cookies stems from the way in which these myriad actors and intermediaries form a large scale network that organizes and stabilizes economic and social relations.

5.2. The cookie, a besieged fortress

The cookie is currently the main standard and the structuring brick of online advertising and marketing. Yet, the actors whose daily activity depends on the proper functioning of cookies have come to evoke the idea of a "post-cookie world" (see for example IAB 2014). How to explain these fears on the part of professionals? The scale of this infrastructure makes it *a priori* almost irreversible, and, from the point of view of the actors, not amendable. Yet, and this is an important lesson from STS, networks and alliances are done and undone. Can digital infrastructures such as those built around cookies de-scale and coalesce as they have scaled?

On closer inspection, the cookie appears to be a fragile compromise. First, the cookie has a number of shortcomings "by design"–which makes sense, as its designers did not anticipate its fortune as a market infrastructure. First, cookies are recorded in web browsers, which can be used by multiple users (family computer) while users can distribute their browsing activity across multiple devices and browsers. In addition, cookies have a limited lifetime. This period is limited by law to 13 months, but many operators operate with cookies that have a lifetime of one month. Finally, cookies can be manually deleted by Internet users in the browser management interface, or automatically when users have installed an ad-blocking extension. These operations reduce the lifespan, coverage, and ultimately the effectiveness of advertising cookies.

The recent evolution of the European regulatory framework is another threat for the cookie. The European General Data Protection Regulation (GDPR), which came into force in May 2018, gives a central place to the free and informed consent of individuals in the process of collection, processing and valuation of personal data. In the areas of marketing and advertising, this is a change with established conventions and practices. The use of advertising cookies is an important part of the debate surrounding the application of this regulation, since the collection of cookie-based personal data will shift from implicit consent after the event (opt-out) to explicit consent beforehand (opt-in). That said, legal uncertainties remain, since the ePrivacy Directive, which is intended to supplement and clarify the application of the GDPR, inter alia for cookies, was still under negotiation in March 2019. In this perspective, the very question of the survival of third-party cookies, for which explicit consent appears difficult to obtain, is posed. As an illustration of this tension, the professional press took hold of this subject and dramatized it by announcing the imminent death of cookies, or even that of targeted advertising.[6]

Following the GDPR, the various players in the value chain have begun to modify their collection mechanisms, or the screening of data collection by their partners, in order to become compliant, or at least to report their efforts towards compliance to regulators. E-mails have been sent to users, information banners have appeared or increased in size on websites, terms of use have been rewritten, etc. Due to the need to revise their personal data policy, companies also resort to consulting firms and external service providers, such as "consent management platforms," in order to ensure their compliance. A few months after the launch of the GDPR, early reports on the consequences of this new regulatory framework on the advertising industry tend to underline that practitioners have become careful on how consent is obtained–especially online commerce and media companies. As a result, middleware companies providing technical means for the advertising industry have set up lists of validated or trusted partners, and thus strengthened their relationship with media websites. This leads, for example, publishers to a thorough examination of partners authorized to deposit third-party cookies on their websites. A study by the Reuters Institute measured a drop of 22% in the average number of third-party cookies found on a sample of news websites in Europe between May and July 2018. This said, the websites reviewed by Reuters still used an average of 60 third-party cookies–most of them being advertising cookies–per webpage after the downturn (eMarketer 2018). In other words, apart from a slight shift in the use of third-party cookies, the implementation of the GDPR has not (yet) questioned the place of the cookie as the central brick of the advertising market infrastructure.

5.3. Deploying the cookie model in a cookie-less world

The development of mobile phone and tablet computer usage represents an additional threat to cookies. In the UK, mobile phones and tablets accounted for 78% of the time spent online by Internet users, and received 68% all digital advertising spending in 2018 (eMarketer). Cookies have developed in desktop web browser environments. Mobile browsing using applications is preferred over browsers like Chrome or Firefox; yet, it is not possible to install cookies in mobile applications. In addition, for Apple smartphones, a large share of web browsing time is spent on the Safari browser, which blocks third-party cookies by default. The incompatibility between cookies and the mobile environment is a serious limitation for advertizers because it raises two types of problems: what identification technology can be used in the mobile environment to replace the cookie? How to "reconcile" computer and mobile usage to keep on tracking users from screen to screen and to create a coherent picture of them?

[6]See for example: "Google is a cookie monster" (eMarketer, August 2018); "European Commission proposal will kill 3rd party cookies" (PageFair, January 2017); "Will 2018 kill user segmentation and targeting (GDPR, Safari browser's restrictions on cookies)?" (Quora, January 2018); "Will GDPR kill the third-party data market?" (MyCustomer, March 2018).

At first glance, although mobile devices are not cookie-friendly, the developers of mobile operating systems have provided the advertising industry with a dedicated substitute for the cookie. On mobile phones, publishers and advertizers can rely on an identifier specifically created for advertising tracking and targeting by mobile operating system developers (Apple for iOS, Google for Android): the IDFA (iOS' Identifier for Advertizers) and Android's Advertising ID. This mechanism makes it possible to match the information collected on a user by different advertising vendors, and to carry out targeting operations from one application to another. In particular, these identifiers can be used to manage targeted and retargeted advertising campaigns on mobile devices. In other words, the market arrangements built around the cookie have set up an industry standard, that has been challenged and stabilized, and that is to be reproduced when the cookie technology is missing. From that point of view, mobile advertising IDs are the testimony of how the cookie infrastructure serves as a model for the digital advertising industry that has to be maintained beyond the cookie itself.

However, this mobile identifier for advertising that copies the cookie is not the center of gravity of the mobile advertising value chain. Mobile advertising amplifies a trend already observed in desktop environments: the rise of "logged" or walled-garden environments, i.e. autonomous milieus built around identifiers specific to an actor, obtained from the connection (log in) to its platform. This strategy is all the more effective if the platform is able to attract a large number of connected users, and to bridge mobile and desktop environments, which is the case of big online services companies such as Google or Facebook. For example, it has been gradually implemented by Facebook under the label "people-based marketing," based on the individual Facebook identifier, active once the user is connected. Facebook is able to track its audience on both its desktop and mobile platforms–where users are permanently logged in. Facebook does not give up using cookies, which serve to track its users outside the platform, *via* sharing buttons, and also to trace users who are not logged in. But cookies are rather a complementary tracking technology for Facebook. In addition, the Facebook advertising platform does not function as an ad exchange that would allow buyers to connect and bid based on information collected from cookies collected from third parties such as data providers. With the exception of the first-party data provided by advertizers (under the "custom audience" label), only Facebook data are operational in the Facebook universe. These data are harvested, gathered and made available to advertizers in the form of targeting tools directly implementable in Facebook's advertising purchase console.

How, from a market infrastructure perspective, can we interpret this fragmentation and the emergence of walled-garden environments that compete with the open infrastructure of cookies? On the one hand, Facebook intends to build an alternative infrastructure, hermetically separated from the cookie infrastructure. In other words, Facebook internalizes knowledge production, capitalization and coordination operations. This infrastructure has become an easily accessible resource for advertizers of all sizes, small and large. Together, they spent more than a quarter of their online advertising investments (desktop and mobile) on Facebook in the UK in 2018, according to eMarketer. But can we still talk about a market infrastructure in the case of Facebook? By internalizing and privatizing critical market operations, Facebook effectively excludes a large number of market participants (with the exception of advertizers, and a few duly accredited partners). In addition, these knowledge production, capitalization and coordination operations are carried out in the background and largely opaque to advertizers, who can only benefit from them by connecting to the one-stop pre-configured and highly automated Facebook advertising console. In other words, advertizers do not access a data infrastructure but Facebook's advertising platform.

In this regards, the Facebook case illustrates the connection between the concept of market infrastructure and that of platform. Plantin et al. (2018) have already explored this connection, and shown how both concepts are relevant to describe digital players such as Facebook or Google. An important part of platforms' activity consists in enrolling and domesticating their environment, in "making things platform-ready" (Helmond 2015), ultimately weakening the players they connect (Pasquale 2015). Mackenzie (2018) analyzes these aspects of platformization as a process of capitalization

and assetization. The Facebook case matches these descriptions: Facebook has used the cookie-based market infrastructure to, little by little, set up an advertising platform, i.e. a closed consumer data environment from which it can extract the most value. In this regards, platforms could be defined as the monopolistic form of market infrastructures, and the Facebook consumer data platform appear as a main pretender to the succession of the cookie-based infrastructure.

6. Conclusion

In this article, we were interested in how the cookie has become the basic building block of a vast market infrastructure that today attracts a significant share of the activities and expenditures devoted to online advertising and marketing. Following an STS approach, we have described how it has gradually been deployed by the advertising industry as a tool for knowledge production, capitalization and market coordination. The cookie has become a standard for the online advertising market. This new mean to engage with consumers through data has deeply modified how marketing professionals operate in digitalized environments. Finally, we have questioned the sustainability of this infrastructure. Growth in mobile usage, regulatory challenges and competing infrastructures will certainly help in the end reduce the centrality and the use of the cookie. On the one hand, the personal data market infrastructure based on the cookie serves as a model and as a mold for complementary or competing infrastructures that use other types of identifiers, such as mobile advertising identifiers. Using these identifiers, ad companies, publishers and advertizers tend to reproduce the main operations permitted by the cookie infrastructure in order to capitalize on their knowledge or clues on consumer preferences and purchase intents. On the other hand, though, there is a tension between the model of a (relatively) open, cookie-based infrastructure and privatized forms of personal data management, based on login credentials. The outcome of this battle of identifiers is uncertain, but it could *in fine* defeat the cookie as an infrastructure simultaneously supporting knowledge production, capitalization and coordination operations.

Based on the study of cookies from a market infrastructure perspective, our theoretical contribution to market studies is twofold. First, we contribute to the framing of the concept of market infrastructure. We identify three operations that such infrastructures support: knowledge production, capitalization, and coordination. Focusing on these three operations appears as an important way of articulating market infrastructures with market devices. Such question is important for market studies, since the concept of market device is central; Chakrabarti et al. (2016) as well as Kjellberg, Hagberg, and Cochoy (2019) pay a particular attention to this articulation. For Chakrabarti et al., the concept of device is not adapted to describing the temporal dynamics of markets, and "moving beyond devices to [market infrastructures], we move beyond snapshots alone which are narrowing and open up to potentially studying the interactions between a variety of devices and their interfaces with [market infrastructures]." Kjellberg et al. go further into this argument: for them, the concept of market infrastructure has a larger scope than market device, and "helps moving beyond mere case studies focused on single and isolated innovations." They conceptualize market infrastructures as some sort of super-devices, "the resulting effect of a successful articulation of various and originally independent market devices" that has faded in the background and "silently supports and structures the consummation of market exchanges." Drawing on the example of cookies, we would indeed agree with the "super-device" understanding of infrastructures, but not as a combination of independent devices. Rather, an infrastructure is a combination of features, a single device supporting market operations that are otherwise distributed among multiple devices. Here comes the importance of temporal dynamics of infrastructures: embedding in a single apparatus the main operations of knowledge production, capitalization, and coordination, appears as the fortuitous result of historical conditions, and as a step-by-step accumulation. In the case of cookies, actors of the nascent advertising market took the opportunity offered by this versatile mechanism initially designed to fill baskets and handle web sessions. Once they are overcharged with key

market-related operations, market infrastructures are like irreformable black boxes on which markets operate.

Second, our work contributes to the study of the digitalization of markets. We have described in detail how the digitalization of the marketing industry has built, upon a simple software mechanism, an effective data infrastructure. What role does the building of such computerized knowledge infrastructure play in the digitalization of markets? In other words, is digitalization a datafication of markets? Our inquiry provides some answers. First, the cookification process is not limited to the advertising industry; it is expanding to other parts of the firms' business activities. Cookie-based consumer data serve today as a connector for previously separate areas: customer relationship management, product design, advertising, store management. The cookie has served as a starting point for the implementation of "data lakes" by advertizers: large repositories aimed at gathering all the data from interactions with consumers, including identifiers, contact information, socio-demographical and behavioral attributes, transaction history, customer service interactions, but also "social data" (consumer reviews, social media data, etc.) and other data sources obtained from third parties such as data brokers. It is in the field of mobile telephony and physical store geography that we situate the new horizon of this cookification movement (Hagberg, Jonsson, and Egels-Zandén 2017; Turow 2017). Such efforts to build data lakes or to digitalize the point of sale are the true legacy of the cookie-based knowledge infrastructure. Another element of the data-based digitalization of markets relates to the ongoing process of assetization of data, and especially consumer data (Beauvisage and Mellet forthcoming). Often described as "new oil," personal data are commodified and traded on the advertising market in the form of contacts, segments or attributes, but they also constitute a specific class of assets for firms when optimizing internal operations or making deals with other firms. In that sense, the datafication of markets looks like a key element of their digitalization.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Thomas Beauvisage is a sociologist and web scientist at the Social Sciences Department of Orange Labs (Sense). His early works and PhD focused on web usage mining and browsing behavior characterization. Today, his activities involve both internet research and market studies; his current research topics cover online participatory market devices, advertising and e-reputation, and uses of online media. He is also involved in methodological investigations on the use of quantitative behavioral material for social science.

Kevin Mellet is a sociologist at the Social Sciences Department of Orange Labs (Sense) and an associate researcher to the LISIS (University Paris-Est Marnes la Vallée). His research is based primarily on economic sociology and science and technology studies and focuses on digital markets, from the standpoint of market intermediaries, businesses, and consumers. He is interested in issues related to online visibility and reputation. More recently, he has been investigating the infrastructures, devices, and practices of the emerging "data marketing" landscape.

References

Ariztia, Tomas. 2015. "Unpacking Insight: How Consumers are Qualified by Advertising Agencies." *Journal of Consumer Culture* 15 (2): 143–162.

Arrow, Kenneth J. 1974. *The Limits of Organization*. New York: WW Norton & Company.

Beauvisage, Thomas, and Kevin Mellet. forthcoming. "Datassets: Assetizing and Marketizing Personal Data." In *Turning Things into Assets*, edited by Kean Birch and Fabian Muniesa. Cambridge: MIT Press.

Bergemann, Dirk, and Alessandro Bonatti. 2015. "Selling Cookies." *American Economic Journal: Microeconomics* 7 (3): 259–294.

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

Beuscart, Jean-Samuel, Kevin Mellet, and Marie Trespeuch. 2016. "Reactivity Without Legitimacy? Online Consumer Reviews in the Restaurant Industry." *Journal of Cultural Economy* 9 (5): 458–475.

Binns, Reuben, Jun Zhao, Max Van Kleek, and Nigel Shadbolt. 2018. "Measuring Third Party Tracker Power Across Web and Mobile." ArXiv:1802.02507 [Cs], February.
Birch, Kean, and Fabian Muniesa, eds. forthcoming. *Turning Things into Assets*. Cambridge: MIT Press.
Bouk, Dan. 2017. "The History and Political Economy of Personal Data Over the Last Two Centuries in Three Acts." *Osiris* 32 (1): 85–106.
Bowker, Geoffrey C., Karen Baker, Florence Millerand, and David Ribes. 2009. "Toward Information Infrastructure Studies: Ways of Knowing in a Networked Environment." *International Handbook of Internet Research*: 97–117.
Bowker, Geoffrey C., and Susan Leigh Star. 1999. *Sorting Things Out: Classification and Its Consequences*. Cambridge: MIT press.
Brookman, Justin, Phoebe Rouge, Aaron Alva, and Christina Yeung. 2017. "Cross-Device Tracking: Measurement and Disclosures." *Proceedings on Privacy Enhancing Technologies* 2017 (2): 133–148.
Brunsson, Nils, and Bengt Jacobsson. 2000. *A World of Standards*. Oxford: Oxford University Press.
Callon, Michel. 1984. "Some Elements of a Sociology of Translation: Domestication of the Scallops and the Fishermen of St Brieuc Bay." *The Sociological Review* 32 (Suppl. 1): 196–233.
Callon, Michel, ed. 1998. *The Laws of the Markets*. Vol. 6. Oxford: Blackwell.
Callon, Michel. 2016. "Revisiting Marketization: From Interface-Markets to Market-Agencements." *Consumption Markets & Culture* 19 (1): 17–37.
Castelluccia, Claude, Stéphane Grumbach, and Lukasz Olejnik. 2013. "Data Harvesting 2.0: From the Visible to the Invisible Web." The Twelfth Workshop on the Economics of Information Security, Allan Friedman, Jun 2013, Washington, DC, United States.
Chakrabarti, Ronika, John Finch, Hans Kjellberg, Clara My Lernborg, and Neil Pollock. 2016. "From Market Devices to Market Infrastructures." Interdisciplinary Market Studies Workshop, St. Andrews, Scotland.
Christl, Wolfie, and Sarah Spiekermann. 2016. *Networks of Control. A Report on Corporate Surveillance, Digital Tracking, Big Data & Privacy*. Vienna: Facultas.
Cluley, Robert. 2018. "The Construction of Marketing Measures: The Case of Viewability." *Marketing Theory* 18 (3): 287–305.
Cluley, Robert, and Stephen D. Brown. 2015. "The Dividualised Consumer: Sketching the New Mask of the Consumer." *Journal of Marketing Management* 31 (1–2): 107–122.
Cochoy, Franck, Pascale Trompette, and Luis Araujo. 2016. "From Market Agencements to Market Agencing: An Introduction." *Consumption Markets & Culture* 19 (1): 3–16.
Ebeling, Mary. 2016. *Healthcare and Big Data*. New York: Palgrave Macmillan US.
Edwards, Paul. 1998. "Y2K: Millennial Reflections on Computers as Infrastructure." *History and Technology* 15 (1–2): 7–29.
Edwards, Paul, Steven Jackson, Melissa Chalmers, Geoffrey Bowker, Christine Borgman, David Ribes, Matt Burton, and Scout Calvert. 2013. *Knowledge Infrastructures: Intellectual Frameworks and Research Challenges*. Ann Arbor: Deep Blue.
eMarketer. 2018. "The GDPR Is a Cookie Monster. Publishers Purge Ad Trackers in Light of Data Regulation." Report.
Englehardt, Steven, and Arvind Narayanan. 2016. "Online Tracking: A 1-Million-Site Measurement and Analysis." In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security - CCS'16*, 1388–1401. Vienna, Austria: ACM Press.
Falahrastegar, Marjan, Hamed Haddadi, Steve Uhlig, and Richard Mortier. 2014. "Anatomy of the Third-Party Web Tracking Ecosystem." ArXiv:1409.1066 [Cs], September.
Favereau, Olivier, and Emmanuel Lazega, eds. 2002. *Conventions and Structures in Economic Organization: Markets, Networks and Hierarchies*. Cheltenham: Elgar.
Federal Trade Commission. 2014. *Data Brokers: A Call For Transparency and Accountability: A Report of the Federal Trade Commission*. Washington, DC: Federal Trade Commission.
Fourcade, Marion, and Kieran Healy. 2017. "Seeing Like a Market." *Socio-Economic Review* 15 (1): 9–29.
Geiger, Susi, Hans Kjellberg, and Robert Spencer. 2012. "Shaping Exchanges, Building Markets." *Consumption Markets & Culture* 15 (2): 133–147.
Ghosh, Arpita, Mohammad Mahdian, R. Preston McAfee, and Sergei Vassilvitskii. 2015. "To Match or Not to Match: Economics of Cookie Matching in Online Advertising." *ACM Transactions on Economics and Computation (TEAC)* 3 (2): 12.
Greenstein, Shane. 2015. *How the Internet Became Commercial: Innovation, Privatization, and the Birth of a New Network*. Princeton: Princeton University Press.
Hagberg, Johan, Anna Jonsson, and Niklas Egels-Zandén. 2017. "Retail Digitalization: Implications for Physical Stores." *Journal of Retailing and Consumer Services* 39 (November): 264–269.
Helmond, Anne. 2015. "The Platformization of the Web: Making Web Data Platform Ready." *Social Media + Society* 1 (2). doi:10.1177/2056305115603080.
IAB. 2014. "Privacy and Tracking in a Post-Cookie World." White paper.
Karpik, Lucien. 2010. *The Economics of Singularities*. Princeton: Princeton University Press.

Kitchin, Rob. 2014. *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*. London: Sage.
Kjellberg, Hans, Johan Hagberg, and Franck Cochoy. 2019. "Enacting a Digital Market Infrastructure in the US Grocery Retail Sector, 1967–2010". In *Thinking Infrastructures: Research in the Sociology of Organizations*, edited by M. Kornberger, G. Bowker, N. Pollock, P. Miller, A. Mennicken, J. Randa Nucho & J. Elyachar, 207–232. Bingley: Emerald Publishing.
Krishnamurthy, Balachander, and Craig E. Wills. 2009. "On the Leakage of Personally Identifiable Information via Online Social Networks." In *Proceedings of the 2nd ACM Workshop on Online Social Networks*, 7–12. Barcelona: ACM.
Larkin, Brian. 2013. "The Politics and Poetics of Infrastructure." *Annual Review of Anthropology* 42 (1): 327–343.
Law, John. 1990. "Introduction: Monsters, Machines and Sociotechnical Relations." *The Sociological Review* 38 (S1): 1–23.
Lyon, David. 2001. *Surveillance Society: Monitoring Everyday Life*. London: McGraw-Hill Education (UK).
MacBeth, Sam. 2017. "Tracking the Trackers: Analysing the Global Tracking Landscape with GhostRank." Technical Report.
Mackenzie, Adrian. 2018. "48 Million Configurations and Counting: Platform Numbers and Their Capitalization." *Journal of Cultural Economy* 11 (1): 36–53.
Mayer, Jonathan R., and John C. Mitchell. 2012. "Third-Party Web Tracking: Policy and Technology." In *2012 IEEE Symposium on Security and Privacy*, 413–427. San Fransisco, CA: IEEE.
Mellet, Kevin, Thomas Beauvisage, Jean-Samuel Beuscart, and Marie Trespeuch. 2014. "A 'Democratization' of Markets? Online Consumer Reviews in the Restaurant Industry." *Valuation Studies* 2 (1): 5–41.
Mills, Jon L. 2008. *Privacy: The Lost Right*. Oxford: Oxford University Press.
Muniesa, Fabian, Liliana Doganova, Horacio Ortiz, Álvaro Pina-Stranger, Florence Paterson, Alaric Bourgoin, Véra Ehrenstein, Pierre-André Juven, David Pontille, and Basak Saraç-Lesavre. 2017. *Capitalization: A Cultural Guide*. Paris: Presses des Mines.
Muniesa, Fabian, Yuval Millo, and Fabian Muniesa. 2007. "An Introduction to Market Devices." *The Sociological Review, Sociological Review Monograph Series* 1–12.
Napoli, P. M. 2003. *Audience economics: Media institutions and the audience marketplace*. New York: Columbia University Press.
Neyland, Daniel, and Sveta Milyaeva. 2016. "The Entangling of Problems, Solutions and Markets: On Building a Market for Privacy." *Science as Culture* 25 (3): 305–326.
Olejnik, Lukasz, Minh-Dung Tran, and Claude Castelluccia. 2014. "Selling off Privacy at Auction." In *Proceedings 2014 Network and Distributed System Security Symposium*, 1–15. San Diego, CA: Internet Society.
Park, J. S., and R. Sandhu. 2000. "Secure Cookies on the Web." *IEEE Internet Computing* 4 (4): 36–44.
Pasquale, Frank. 2015. *The Black Box Society: The Secret Algorithms That Control Money and Information*. Cambridge: Harvard University Press.
Peacock, Sylvia E. 2014. "How Web Tracking Changes User Agency in the Age of Big Data: The Used User." *Big Data & Society* 1 (2): 205395171456422.
Plantin, Jean-Christophe, Carl Lagoze, Paul N Edwards, and Christian Sandvig. 2018. "Infrastructure Studies Meet Platform Studies in the Age of Google and Facebook." *New Media & Society* 20 (1): 293–310.
Schneier, Bruce. 2015. *Data and Goliath: The Hidden Battles to Collect Your Data and Control Your World*. New York: WW Norton & Company.
Schwartz, John. 2001. "Giving the Web a Memory Cost Its Users Privacy." *New York Times* 4 (1): Section A, Page 1.
Siebecker, Michael R. 2002. "Cookies and the Common Law: Are Interest Advertisers Trespassing on Our Computers." *Southern California Law Review* 76: 893.
Simon, Herbert A. 1991. "Organizations and Markets." *Journal of Economic Perspectives* 5 (2): 25–44.
Smith, Mike. 2014. *Targeted: How Technology Is Revolutionizing Advertising and the Way Companies Reach Consumers*. New York: Amacom.
Star, Susan Leigh. 1999. "The Ethnography of Infrastructure." *American Behavioral Scientist* 43 (3): 377–391.
Star, Susan Leigh, and Geoffrey C. Bowker. 2002. "How to Infrastructure." In *Handbook of New Media: Social Shaping and Consequences of ICTs*, edited by Leah Lievrouw and Sonia Livingstone, 151–162. London: SAGE.
Star, Susan Leigh, and Karen Ruhleder. 1996. "Steps Toward an Ecology of Infrastructure: Design and Access for Large Information Spaces." *information Systems Research* 7 (1): 25.
Turow, Joseph. 2012. *The Daily You: How the New Advertising Industry Is Defining Your Identity and Your Worth*. New Haven: Yale University Press.
Turow, Joseph. 2017. *The Aisles Have Eyes: How Retailers Track Your Shopping, Strip Your Privacy, and Define Your Power*. New Haven: Yale University Press.
von Hayek, Friedrich August. 1937. "Economics and Knowledge." *Economica* 4 (13): 33–54.
von Hayek, Friedrich August. 1945. "The Use of Knowledge in Society." *The American Economic Review* 35 (4): 519–530.

Yu, Zhonghao, Sam Macbeth, Konark Modi, and Josep M. Pujol. 2016. "Tracking the Trackers." In *Proceedings of the 25th International Conference on World Wide Web - WWW '16*, 121–132. Montréal: ACM Press.
Zimmer, Michael. 2008. "The Gaze of the Perfect Search Engine: Google as an Infrastructure of Dataveillance." In *Web Search*. Vol. 14., edited by Michael Zimmer, 77–99. Berlin, Heidelberg: Springer.
Zuboff, Shoshana. 2019. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. London: Public Affairs.