Regulation of data-driven marketing and management theory: bibliometric analysis, systematic literature review and research agenda

Jorge Xavier and Winnie Ng Picoto
Department of Management, ISEG Lisbon School of Economics and Management, Lisboa, Portugal

Abstract

Purpose – Regulatory initiatives and related technological shifts have been imposing restrictions on data-driven marketing (DDM) practices. This paper aims to find the main restrictions for DDM and the key management theories applied to investigate the consequences of these restrictions.

Design/methodology/approach – The authors conducted a unified bibliometric analysis with 104 publications retrieved from both Scopus and Web of Science, followed by a qualitative, in-depth systematic literature review to identify the management theories in literature and inform a research agenda.

Findings – The fragmentation of the research outcomes was overcome by the identification of 3 main clusters and 11 management theories that structured 18 questions for future research.

Originality/value – To the best of the authors’ knowledge, this paper sets for the first time a frontier between almost three decades where DDM evolved with no significative restrictions, grounded on innovations and market autoregulation, and an era where data privacy, anti-trust and competition and data sovereignty regulations converge to impose structural changes, requiring scholars and practitioners to rethink the roles of data at the strategic level of the firm.

Keywords Data-driven marketing, Management theory, Regulation, Martech, Adtech

Paper type Literature review

1. Introduction

To the best of our knowledge, the idea that data is the new oil was stated for the first time by Clive Humby in 2006, at ANA Senior Marketer’s Summit, Kellogg School (Mazurek and Malagocka, 2019; Palmer, 2006). Companies that master data-driven processes exhibit higher levels of customer engagement (Grandhi et al., 2020) and the number of use cases for marketers has considerably increased, raising new requirements, competencies and skills in domains that before were reserved for data science professionals (Saura et al., 2021). The academy has been studying different perspectives of the data-driven marketing (DDM) field such as conceptual frameworks to leverage customer data (Kumar et al., 2013), the impact of...
the evolving technological landscape on the role of marketers (Quinn et al., 2016) or the new
talent gaps and organizational design arrangements (Leeflang et al., 2014). However, as
Palmer wrote, “data is just like crude. It’s valuable, but if unrefined it cannot really be used”
(Palmer, 2006). DDM requires investment in resources, infrastructure, systems and expertise
(Grandhi et al., 2020). Data must be created, structured, stored and maintained. To perform
these processes, marketers are confronted with technical, ethical and political choices
(Cluley, 2020). In the case of advertising, brands often subcontract agencies to run the
campaigns, following practices that evolved together with the internet since 1994 (Taylor,
2009) and being highly dependent on third-party data. Most of the data that feeds the
digital advertising ecosystem still relies on cookies placed on websites by third-party data brokers
(Neumann et al., 2019). The digital consumer profiles are then sold to marketers as
predefined audiences for targeting (Mellet and Beauvisage, 2020; Neumann et al., 2019).
Recent changes in regulation have been imposing increasing challenges on the use of
personal data for marketing purposes. In 2018, the General Data Protection Regulation (GDPR)
came to force (European Commission, 2016), imposing principles to protect individuals’
personal information and privacy. Regulations such as the GDPR, the Australian Privacy Act or the
California Consumer Privacy Act aim to protect consumers but raise new challenges and costs
for companies to comply with those regulations (Quach et al., 2022). Similar initiatives have
been emerging around the world. For example, the General Law for Data Protection in Brazil,
the Personal Data Protection Act in Thailand, the Draft Personal Data Protection Bill in India
or the Draft of Law on Protection of Personal Information in China (Hsu, 2021). In January 2020,
Google announced plans to withdraw support for third-party cookies in Chrome (Google, 2020).
This process was delayed twice. First Google postponed the phase-out of third-party cookies to
the end of 2023 (Google, 2021). Then, a new announcement postponed it to the end of 2024
(Google, 2022).

The impact for the organizations has been tremendous (Latvala et al., 2022). Privacy
concerns, the role of data privacy in marketing and the effects on firm performance have
been studied by leading authors (Martin et al., 2017; Martin and Murphy, 2017; Okazaki
et al., 2020). Beyond privacy, public authorities seek further control to guarantee data
sovereignty through massive regulation and the creation of data spaces (European
Commission, 2020a). Firms’ data-driven activities go far beyond marketing, but arguably, no
other business function is being more impacted, because most of the restrictions fall on the
personal data used by marketers for personalization purposes. The availability of data and
the capacity to analyse it are among the key objectives for relational marketing capabilities
to drive Customer Relationship Management systems adoption (Pedron et al., 2016). The
existing literature reviews do not address how these restrictions on data usage are affecting
management activities. For example, the literature review from Saura (2021) follows an
operational perspective, including studies from Computer Science, Information Systems and
Engineering domains. The review from Fernández-Rovira et al. (2021) addresses digital
transformation in general and the use of big data. Similar studies (Akter et al., 2021) explore
big data to enhance firms’ growth. Our study seeks a better understanding of the
implications for management in result of the increasing restrictions on DDM.

2. Method

Bibliometric methods can be an advantage in literature reviews, before the analysis of the
documents, guiding the researcher through the relevant works and mapping research fields while
reducing subjectivity (Zupic and Cater, 2015). In this paper, we combine a systematic review with
a bibliometric analysis, integrating the quantitative and qualitative efforts, after noticing that it
has been frequent across different domains, such as financial studies (Ansari et al., 2022),
2.1 Study design
In the research design, we declared the two questions that are key for our research:

RQ1. Which are the main restrictions on data-driven marketing?

RQ2. Which are the key management theories applied to investigate the consequences of these restrictions?

To build the database of articles to address our research questions, we adopted PRISMA protocol for systematic reviews, described in Figure 1. This approach was developed in health-care research communities (Moher et al., 2009) but has been increasingly used for systematic reviews in other research fields, such as digital marketing (Saura, 2021).

2.2 Data collection
For the data collection, we started with a broad search of the Clarivate Web of Science database. We used the following string for the search: (TS = Online Advertising OR TS = Behavioral Advertising OR TS = Digital Marketing) AND AK = data AND (WC = Business or WC = Management). Additionally, author keywords should include “data” and the Web of Science Category (WC) is set to be “Business” or “Management”. Then, the query was refined with two additional filters: document types = “Articles” and research areas = “Business Economics.” No additional filters were applied. The final query was executed in 2022, May 8th, resulting in 185 articles, as described in Figure 1. After analysis of the titles, abstracts and keywords, we removed the articles not related to research questions, not relevant to research areas, with a total of 104 articles. The themetic map technique included all the articles, organized like this: cluster 1, 29 articles; cluster 2, 30 articles; cluster 3, 37 articles. The 104 articles were then included in the Systematic Literature Review (SLR).

Figure 1. PRISMA flow diagram for systematic literature review (Moher et al., 2009)
goals and not conclusive (Moher et al., 2009; Zupic and Čater, 2015). At the end of this process, 72 articles were selected. Then, we applied the same search procedures to the Scopus database to assure completeness and robustness to our sample, in line with the recommendations in literature to use both the Scopus database and Web of Science in bibliometric analyses (Echchakoui, 2020). The two databases complement each other. By using both, we can get richer research (Sánchez et al., 2017). The results from Scopus allowed us to get 79 articles, the last update on May 15th, 2022. We followed the same qualitative screening protocol to select the articles to include in the bibliometric analysis, and 49 articles were considered relevant to the research goals.

2.3 Bibliometric analysis
We followed the procedures described in the literature to merge the results from both databases and conduct a unique integrated analysis. Based on the works of Echchakoui (2020) and Caputo and Kargina (2022), we used the code available on R documentation to merge the two databases and then remove the duplicated articles (Aria and Cuccurullo, 2017; Caputo and Kargina, 2022). The merged database included 104 articles. Although we did not apply any filter regarding publication dates, the time span of the articles is the period between 2003 and 2022 as described in Figure 2, which was expectable because of the novelty of the studied topic. The cluster analysis based on similarities (Zupic and Čater, 2015) was the data analysis technique selected to analyse the 104 papers that passed on the screening process. The open-source software R Studio with Bibliometrix package was used for the cluster analysis, also known as Coupling Map (Aria and Cuccurullo, 2017).

3. Results
3.1 Cluster analysis
We analyzed the knowledge structures to understand the emergent concepts associated with data as a valuable and regulated resource and how those emergent concepts are related to each other (Ahuja and Novelli, 2015; Aria and Cuccurullo, 2017; Jonassen and Wang, 1992). The co-word analysis (Callon et al., 1983) uses the most relevant words of documents to examine the theoretical structure of a research field. We selected the abstracts as coupling

![Figure 2. Depicts that the past five years are accountable for more than half of all the selected publications in this field](image)

Source: Figure by authors
parameters to address research questions searching for the main topics associated with the research goals (Zupic and Cater, 2015). Our clustering data analysis relied on the Walktrap community detection algorithm, proposed by Pons and Latapy (Pons and Latapy, 2006):

\[ \Delta \sigma(C_1, C_2) = \frac{1}{n} \left( \sum_{i \in C_3} r^2 iC_3 - \sum_{i \in C_1} r^2 iC_1 - \sum_{i \in C_2} r^2 iC_2 \right) \]

Walktrap algorithm has been highly recommended (Gates et al., 2016; N.R. Smith et al., 2021) and works well with graph sizes as small as 100 studies. This was decisive for our work because the data set contained 104 studies. For each pair of adjacent communities \{C_1,C_2\}, the variation \( \Delta \sigma(C_1, C_2) \) of \( \sigma \) that would be generated if \( C_1 \) and \( C_2 \) are merged into a new community \( C_3 = C_1 \cup C_2 \) is computed. The quantity would then depend on the vertices of \( C_1 \) and \( C_2 \), not on the further communities. Then the two communities that give the minimum value of \( \Delta \sigma \) are merged.

The result was the coupling map with three clear clusters described in Figure 4, revealing the strength of association between information items (Callon et al., 1983) with clusters representing conceptual groups of topics or themes studied in a certain research field (Cobo et al., 2011b).

3.1.1 Cluster 1 – the balance of welfare and privacy. Cluster 1 is represented on the coupling map in Figure 4, at the left quadrants and the following terms on the label: personal information; display advertising; data protection; online advertising; and personal data. As described in Table 1, the centrality value of 0.92 and an impact value of 2.07 suggest that those studies have some focus but do not provide a major contribution to the research field. Cluster 1 includes 29 articles, the older one from 2003 and all the others since 2010. The most recent article was published in 2022. This cluster explores the importance of data and the existing conflictual interests. On one side, the informative role of advertising and the better experience provided for consumers and, on the other side, the risk of privacy concerns (Tucker, 2012). Some authors argue that the advertising funding supports much of the free content available on the internet, that otherwise would be payable to get access to it (Christiansen, 2011). This conflictual interest in privacy and the communities’ desire to
engage digitally have been relevant for the new concepts of data capitalism (West, 2019). The understanding of some authors is that companies and consumers choose to share data because it is useful to both (Mazurek and Malagocka, 2019). The research included in this cluster pointed out that the use of personal data allows the customization of products and services tailored to the preferences of each customer, as well as price discrimination that can be beneficial both for buyers and sellers (Krämer et al., 2019; Rayna et al., 2015). Nevertheless, recent studies included in this cluster conducted experiments using ad-blockers applications but did not find evidence that the blocking of targeted advertising affects consumer welfare, either on prices paid, search costs or even product satisfaction, challenging the trade-off of higher privacy in exchange for consumer welfare (Todri, 2022). Moreover, the level of knowledge and consciousness from consumers about their personal data dictates significative differences in the willingness of consumers to pay to keep their privacy (Li and Nill, 2020). The studies in this cluster since 2010 claim the need for legal frameworks to improve protection over consumers’ privacy and personal data, self-regulation and privacy-friendly technologies (King and Jessen, 2010; Mathews-Hunt, 2016). Much of the studies in this cluster are dedicated to the balance between welfare and privacy. However, they are not the key contributors from the perspective of the management and the firm. The literature review revealed that none of these studies were grounded on an identifiable management theory, and therefore, those studies were not selected for the discussion in Table 2.

---

**Table 1. Clustering by coupling resume**

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Frequency</th>
<th>Centrality</th>
<th>Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>29</td>
<td>0.92</td>
<td>2.07</td>
</tr>
<tr>
<td>2</td>
<td>38</td>
<td>1.01</td>
<td>2.71</td>
</tr>
<tr>
<td>3</td>
<td>37</td>
<td>0.94</td>
<td>1.73</td>
</tr>
</tbody>
</table>

*Source: Table by authors*
<table>
<thead>
<tr>
<th>Cluster</th>
<th>Theory</th>
<th>Contributions</th>
<th>Restrictions</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Adaptive capabilities</td>
<td>Gap between the organization capabilities and the complexity raised with big data, media spread, more touchpoints and the segment-of-one (Day, 2011)</td>
<td>The digital marketing capabilities gap not only requires adaptive capabilities on data management but also affects decisions about internal development versus outsourcing of specialized services</td>
</tr>
<tr>
<td>2</td>
<td>Affordance theory</td>
<td>The investment (in big data) fails to achieve the expected results. Firms need new approaches and tools to address big data payoff (De Luca et al., 2021)</td>
<td>The accountability of investments in big data emerges as a signal of the firm’s maturity. Data-lakes with large volumes of data increase the costs and risks for the firms and do not pay-off unless the data is actionable and oriented to business results. Data is not a subproduct of IT departments and should be managed at the level of strategic management. In the long term planning and data-governance orientations, the management should work to overcome the restrictions on the value extraction for citizens and businesses, affecting consumers’ privacy, public sector transparency and market power balance.</td>
</tr>
<tr>
<td>2</td>
<td>Business ethics</td>
<td>Respect for individual rights and safeguard of ethical principles on digital interactions with customers and personal data management (Fernández-Rovira et al., 2021)</td>
<td>The management should revisit long-term partnerships with the largest platforms, also known as gatekeepers, incorporating the thoughts on literature from business models of two-sided markets structures. Regulatory efforts to limit the market power of the largest platforms such as the Digital Services Act and the Digital Markets Act (European Commission, 2020b) might limit the extent of what platforms will be able to offer in the future.</td>
</tr>
<tr>
<td>3</td>
<td>Business models</td>
<td>Uncertainty on the trade-off between personal data and free services to finance two-sided markets structures that support advertising (Trabucchi et al., 2017)</td>
<td>The management should revisit long-term partnerships with the largest platforms, also known as gatekeepers, incorporating the thoughts on literature from business models of two-sided markets structures. Regulatory efforts to limit the market power of the largest platforms such as the Digital Services Act and the Digital Markets Act (European Commission, 2020b) might limit the extent of what platforms will be able to offer in the future.</td>
</tr>
<tr>
<td>2</td>
<td>Diffusion of innovation</td>
<td>Managers’ willingness to understand the challenges of DDM and explore its potential to improve business results (Grandhi et al., 2020)</td>
<td>The firms’ investment in resources in the field of DDM is recognized as a competitive factor and a way to lead the perception of consumer behaviors. It should be driven by top management within the organizations, as it might increase the differences between early adopters and laggards.</td>
</tr>
<tr>
<td>3</td>
<td>Efficient market theory</td>
<td>Difficulty on understanding unstructured data online before product release (Xiong and Bharadwaj, 2014).</td>
<td>Data-driven applications can provide insights into consumer sentiment without compromising their privacy. However, more capabilities are needed to explore unstructured online data</td>
</tr>
<tr>
<td>Cluster</td>
<td>Theory</td>
<td>Contributions</td>
<td>Restrictions</td>
</tr>
<tr>
<td>---------</td>
<td>--------</td>
<td>---------------</td>
<td>--------------</td>
</tr>
<tr>
<td>3</td>
<td>Psychological reactance (PRT) and communication privacy management (CPM) theories</td>
<td>By increasing trust and transparency on consented relationships with consumers, it is possible to mitigate the perception of intrusion and strength customer engagement (Brinson and Britt, 2021)</td>
<td>The consumer’s increasing desire to adopt ad-blockers is a signal of their own will that goes beyond regulation initiatives and should raise further questions such as what these consumers expect, even when interacting with the firm’s own media</td>
</tr>
<tr>
<td>2</td>
<td>Resource-based view (RBV)</td>
<td>Knowledge extracted from DDM activities is considered a key resource, as valuable as oil (Akter et al., 2021)</td>
<td>Big data become prominent in the scope of firm’s tangible and intangible resources and is key for business strategy. Nevertheless, the tension among firms, consumers and regulators is growing. The costs of data breaches increased 13% in the past two years (Ponemon Institute, 2022) and class-action lawsuits are encouraging consumers to be more active about their own rights. The value of big data also depends on a safe and ethical relationship with consumers.</td>
</tr>
<tr>
<td>3</td>
<td>Signaling theory</td>
<td>Inadequate measures of media frequencies for customer journey data (Klein et al., 2020)</td>
<td>The measurement of media frequencies for customer journey data is a challenge for marketers and the ability to measure and report the value of the outcomes will benefit from new metrics of cross-media exposure</td>
</tr>
<tr>
<td>2</td>
<td>Structuration theory</td>
<td>Tensions among firms, consumers and regulators, caused by technologic changes and privacy constraints (Quach et al., 2022)</td>
<td>The strategic orientation based on big-data to drive firms’ performance and expansion into global markets should also consider the differences in the regulatory context of each market because the privacy regulations are not worldwide harmonized</td>
</tr>
<tr>
<td>3</td>
<td>Value theory for personal data</td>
<td>Personal data might be scarcer and more expensive to obtain in the future because consumers impose demanding pricing schemas to share their data (Spiekermann and Korunovska, 2017)</td>
<td>Managers should be aware that personal data might be scarcer and more expensive in the future, even if the increase of digital interactions theoretically allows firms to collect more data. But eventually, it is a natural consequence imposed by regulations and the consciousness of the data subjects about their own rights</td>
</tr>
</tbody>
</table>

Table 2. **Source:** Table by authors
3.1.2 Cluster 2 – data-driven marketing maturity and the challenges for the management. Cluster 2 is represented in Figure 4, at the top-right quadrant, with blue color and the following terms on the label: digital engagement; digital technologies; social media; digital marketing; and machine learning. Cluster 2 has a centrality value of 1.01 and an impact value of 2.71, suggesting that these studies are dedicated to “motor themes” (Cobo et al., 2011a). Therefore, well-developed and relevant themes structure the research field. Cluster 2 includes 38 articles, from 2005 to 2022. The literature in this cluster highlights the importance of DDM for the top management of the firm. Data-driven service marketing has the potential to elevate marketing from a cost center to a value creation center (Kumar et al., 2013), but for the laggards in DDM adoption, this might be a source of competitive disadvantage (Day, 2011). Other contributions are found in studies dedicated to the organization challenges such as gaps in skills and talent, processes and organizational design as well as precise and actionable metrics DDM (Leeflang et al., 2014). These authors also highlight the threat for marketers being replaced in their job functions, by technical and digitally oriented professionals. The studies recommend more training for marketers and managers in disciplines that can improve the capacity to use and get actionable insights from data (Kumar et al., 2013). By exploring some fragilities, other authors point out that, if the marketers can generate tailored approaches to each customer, then it is also true that nearly half of the projects fail to achieve the expected results (Pranjić and Rekettye, 2019).

Some authors (Micheaux and Bosio, 2019; Quinn et al., 2016) argue that although marketing professionals do not need to be data scientists, they need to understand the role and way of thinking about data, to eventually contract or manage them, take decisions and assume data officers roles. New courses have been proposed following a “data as a service” view in the customer journey (Micheaux and Bosio, 2019). Quinn et al. argue that DDM has a substantial impact on marketing practice and marketing function, as a discipline of management, but it been often grabbed in the hands of intermediaries such as digital agencies, with their own commercial agendas (Quinn et al., 2016). This line of research is reinforced in the context of increasing regulation because the bigger reliance on first-party data imposes more complexity on outsourcing activities. When brands, as data controllers, work with third parties (data processors) and share personal data, they are assuming risks and vulnerabilities that must be properly managed (Gregory and Bentall, 2012). From a functional perspective, a data-driven market is often empowered by data sciences to boost digital marketing strategies for sales, generate brand awareness and gain access to new markets (Saura, 2021). But some authors are also studying the technologies and processes to tie digital marketing and supply chain management (Ardito et al., 2019) or even to evolve DDM from the firm’s boundaries to the policy research and the big societal trends, such as poverty, sustainability, health or education (Sheth and Kellstadt, 2021). Big data, hereby understood as the sum of structured and unstructured data that feeds modern DDM, has also been studied in the context of resource-based view (RBV) theory, where knowledge, that is extracted from data, is considered a key resource (Akter et al., 2021) and research included in this cluster shows the strategic pathway in the big data spectrum toward knowledge a valuable resource (Sakas et al., 2021). The firms exploring big data often fail to attain better performance and the promised advantages (De Luca et al., 2021). One of the solutions mentioned by different authors is the machine learning application, to increase the speed and knowledge processing of firms and consumers, because of the capacity to handle data at scale, including unstructured data, though claiming for more transparency and easier interpretation (Ma and Sun, 2020). This line of research recommends that marketers should explore big data to help the firms to reach the right audience, retain customers, avoid churn and grow business and profitability through customer-centricity approaches (Grandhi et al., 2020).
Machine learning techniques can also be effective to predict demand and determine the market opportunity for new products and services across industries (Mariani and Fosso Wamba, 2020).

Finally, cluster 2 also offers studies on firm–consumer interactions fostered by the data privacy requirements, embracing the viewpoints of consumers, regulators and firms (Quach et al., 2022) and the necessity to use ethical principles, respect for individual rights, environment, sustainability and economic welfare in the digital interactions with customers (Fernández-Rovira et al., 2021).

3.1.3 Cluster 3 – business models and the value of personal data. Cluster 3 is represented on the coupling map in Figure 4, at the centre, with green color and the following terms on the label: clickstream data; ad blocker; personalized advertising; banner ads; and personal data. Cluster 3 has a centrality value of 0.94 and an impact value of 1.73, suggesting that these studies are important for the research field but still underdeveloped. Cluster 3 includes 37 articles, and as cluster 2, the articles were published between 2005 and 2022, including contributions to business models theory. Trabucchi et al. (2017) examined the business models of two-sided markets structures, with consumers accessing a service on one side and the firms who pay for the advertising or the access to data on the other side. They indicate that big data has the potential to empower new business models, where users gain access to services for free, in exchange for sharing their valuable data stream. An exploratory study about data sharing on Facebook (Spiekermann and Korunovska, 2017) found that when consumers became aware that their data personal is a commercial resource, they start valuing it much more. Aiolfi et al. (2021) points that firms should invest in the relevance and credibility of their messages but declare that, although people are worried about their privacy, they might not act according to it. Mellet and Beauvisage (2020) provide a relevant contribution on digital marketing infrastructure, supporting the advertising ecosystem.

Other studies suggest that third-party audiences frequently have a poor cost-benefit relation, because of high costs and associated inaccuracy (Neumann et al., 2019). Therefore, the decline of third-party data might be an opportunity for a different data strategy, based on a first-party data paradigm to overcome the existing restrictions and to deliver the benefits of a stronger and relevant customer engagement (Latvala et al., 2022). Another research line is explored by studies dedicated to the way the firms allocate their marketing investments and suggest that spreading the budget across the different types of media improves brand awareness and conversion rates across the customer journey (Klein et al., 2020). This thesis is reinforced by Bharadwaj et al. (2020), who argue that earned media, own media and paid media are complementary and decisive for the consumer decision journey. Contributions that link privacy regulation with DDM practices are also found in cluster 3. In an interpretation exercise, Thomas (2019) selects the two principles from GDPR that have been driving most of the efforts for marketers, namely: the explicit consent from individuals to process their data and the demonstration of a legitimate interest on data processing. A different study dedicated to ad blocking (Brinson and Britt, 2021) followed a different path to elevate the importance of consent, demonstrating how strengthening trust and fostering transparent and consented relationships with consumers may mitigate the perception of intrusion and strengthen the engagement levels.

4. Discussion and proposal of a research agenda
In this section, we selected those studies that are clearly grounded on and related to the management theory. After identifying the restrictions of DDM in a regulated world (RQ1), we used the 11 bodies for knowledge identified and described in Table 2, to discuss the implications to management (RQ2).
4.1 Adaptive capabilities
The research from Day, G. is dedicated to adaptive marketing capabilities, a theory that follows an outside-in orientation (Day, 2011). This means that managerial practices start from the market first, anticipating and reacting to market trends. It differs from the resource-based view and dynamic capabilities, theories that tend to be grounded on an inside-out orientation, from the firm to the market. DDM has been relevant for those companies that by adopting vigilant market practices, open marketing and adaptative market experimentation practices need to process both structured and unstructured data, to get insight and feedback from the market. For future research, it is important to investigate the restrictions on data usage on these three pillars of the adaptative capabilities. Regulation has been imposing barriers on data sharing between companies, forcing the disclosure to the data subjects and clarifying the roles of data controllers and data processors. This impacts the idea of marketing through open networks of partners, such as external agencies, data providers or text miners. As the author anticipated, these networks have been providing value and power to organizations in the past decade. However, this market architecture is now threatened.

4.1.1 Example of questions for future research.
- How can outside-in orientations thrive with data-driven restrictions?
- What new adaptive capabilities should be forged in an increasingly regulated world?

4.2 Affordance theory
Affordance theory emerged from ecological psychology rooted in the thesis that affordance value depends on the observer’s perception (Gibson, 1986). Authors such as Volkoff and Strong (2013) applied the affordance theory to organizations and information systems, highlighting the relation between the IT artifacts and the organizational actors. The studies from De Luca et al. (2021) contributed to materializing big data marketing affordances and amplified the evidence on how big data investments are related to performance. However, because the affordances are dynamic, it changes in response to the relationship between the actors and the environment. Recent regulatory efforts raised uncertainty and amplified the risks for the firms associated with data management. In particular, the personal data used in DDM. New actors such as legal and security experts or data protection officers flourish in the organizations. There are more obstacles for the firm to activate data-lakes not because of technical reasons but because of regulatory constraints. Data investments are elevated to the strategic level of the firm.

4.2.1 Example of questions for future research.
- Is there a moderator role of regulation on the value perception of DDM activities?
- Might regulation be a mediating variable on the relation of big data investments with the firm performance?

4.3 Business ethics
In 1986, Mason (1986) identified four key issues of ethics for the information age, namely: privacy, accuracy, property and accessibility, known as PAPA. The studies from Cazier and LaBrie (2011) investigated the consumer’s perception and reaction to the use of their personal data, finding common positive and negative myths on these topics and evaluating the degree of ethical or unethical practices through consumers’ eyes. The respect for
individual rights and safeguarding of ethical principles on digital interactions with customers and personal data management were topics addressed in the research of Fernández-Rovira et al. (2021). Much of Manson’s thoughts about ethics structured the modern regulatory efforts. Informed on applying the normative theories of business ethics in the firm and information systems domain (H. J. Smith and Hasnas, 1999), it is important that future research examine how the increasing consumers’ consciousness and control over personal data influence business ethics, following the ideas of Cazier and LaBrie, because consumers’ beliefs about the value for individuals, business and society are what ultimately defines what is an ethical or unethical data practice.

4.3.1 Example of questions for future research.

- Might consumers’ data consciousness play a moderator role in the relationship between DDM and business ethics?

4.4 Business models

The trade-off between personal data and free services to finance the two-sided markets structures of digital advertising studied by Trabucchi et al. (2017) is a tribute to the business models that flourished in almost self-regulated ecosystems. Since the publication from Amit and Zott (2001) addressing the value creation in electronic business, the literature on business models has been very rich, fueled by all the transformations leveraged by the internet on the system of activities executed by the firm. Platforms form a specific business model that become prominent in the advertising industry, with its own theoretical body of knowledge (Gawer and Henderson, 2007; Tee and Gawer, 2009). Recent literature suggests that business models constitute a theory, exhibiting empirical confirmation and theoretical justification (Sierotowicz and Sierotowicz, 2018). Future research should address how the two-sided markets structures of advertising might be impacted by regulation. Not only data privacy regulation but also the antitrust and competition regulations limit the market power of largest platforms such as the Digital Services Act and the Digital Markets Act (European Commission, 2020b). These regulatory initiatives include new and specific obligations for online platforms and very large platforms. These raise research questions for the firm and affect decisions on long-term partnerships with the largest platforms also known as gatekeepers. The research from Trabucchi et al. (2017) studied a myriad of applications that collect a broad range of data, either from sensors or user-entered data. The apps from platform providers such as Google or Apple were excluded because those apps are considered to be part of a greater product service system. This dual role of the platforms suggest that they might benefit their own services in unfair conditions with those business dependents on them to reach the consumers.

4.4.1 Example of questions for future research.

- Will regulation generate more challenges to platforms or to the business that rely on them?
- How can the firms that produce non-personal data guarantee some protection over this asset?

4.5 Diffusion of innovation

The research from Grandhi et al. examines DDM practices and the way firms can improve shareholder value through enhanced use of data, following diffusion of innovation theoretical model (Rogers, 2004) and embracing the study of factors that influence the
adopters’ willingness and capacity to foster the change processes (Grandhi et al., 2020). The Diffusion of Innovation theory has been used in the field of marketing, to study the expansion of new products, and in the field of management, to investigate innovation diffusion in the context of the firm. In future research, it would be important to understand how the challenges imposed by regulation on data-driven practices interplay with Rogers’ Diffusion of Innovation.

4.5.1 Example of questions for future research.

- Is regulation reducing the differences between early adopters and laggards?
- Are the factors associated with regulation such as costs, technology, competences and risks increasing the divide between the adapters?

4.6 Efficient market hypothesis

The efficient market hypothesis is a recognized theory in finance, rooted in the idea that the stock market is efficient and prices reflect the information available (Malkiel, 2003). In the field of marketing, it has been used in studies that examined the relationship between the value of firm’s advertising and stock prices (Oh et al., 2016) or the impact of third-party product reviews of firm value (Chen et al., 2012). The research from Xiong and Bharadwaj (2014) examined the impact of online buzz before product release. In this regard, future research should examine how regulation affects the information available and actionable. The requirements to ask consumers’ consent for using their data for analytics purposes, the power of the big platforms that work as gatekeepers in walled gardens of data and the different regulatory frameworks across industries and regions might limit the data available and generate asymmetric information.

4.6.1 Example of questions for future research.

- Can users under different regulatory frameworks generate asymmetric information affecting the pre-release buzz evolution curve?

4.7 Psychological reactance and communication privacy management theories

Psychological reactance is an impulse from people in response to constraints on their liberty, that has been used in marketing literature to study the consumers’ behaviors in domains such as online recommendation services (Lee and Lee, 2009), advertising (Baek and Morimoto, 2012) or loyalty programs (Chang and Wong, 2018). Brinson and Britt (2021) used this body of knowledge extended with concepts from communication privacy management theory, which provides constructs to understand how people form decisions on sharing their private information (Petronio, 2015). Brinson and Britt studied how some antecedents influence consumers’ trust and contribute to consumers’ increasing desire to install ad-blockers. The findings reveal that personalized advertising is associated with more negative outcomes when compared with traditional advertising. However, when skepticism is low and trusts in high, people tend to exhibit a positive attitude about personalized advertising and, therefore, less propensity to install ad-blockers. While uncertainty dominates the advertising ecosystem with the sunset of third-party cookies, future research should go beyond ad-blocking software and investigate how brands can build trust and respect with consumers to personalize not only online ads but also communications within the firm’s own media.

4.7.1 Example of questions for future research.

- Might skepticism and trust play a moderator effect on people’s decisions to share private information on cross-media engagement?
4.8 Resource-based view

The work from Akter et al. contributed to reducing the gap between the strategic orientation based on big data and its impact in the field of international marketing, grounded on the resource-based view body of knowledge (Akter et al., 2021). The RBV theory is of paramount importance for marketing researchers, at least since the works from Srivastava et al., dedicated to the importance of market-based assets for the competitive advantage of the firm, integrating RBV and marketing theories and practice (Srivastava et al., 2001). Built up on RBV, Teece developed the dynamic capabilities theories as the result of the convergence of assets, processes and evolutionary paths (Teece et al., 2009), following the same “inside-out” perspective of RBV, as already mentioned in our work (Day, 2011). The studies from Srivastava et al. already included customer and market data within the market-based assets (Srivastava et al., 2001). Since then, the importance of data for the firm did not stop to growth and data-driven strategic orientation acquired increasing importance. Akter et al. identified three key issues in the field of international marketing, namely: digital platform orientation, market orientation and entrepreneurial orientation relying on big data as a key resource (Akter et al., 2021). However, the extension of data availability, the complexity to handle it and the associated costs are examples of variables that raise further questions.

4.8.1 Example of questions for future research.

- How the confluence of different regulatory initiatives embracing data privacy, antitrust and competition, as well as data sovereignty protection might impact big data, as a key resource for international marketing?

4.9 Signaling theory

Signaling theory analyses behaviors when two parties have access to different information (Connelly et al., 2011). Marketing researchers adopted this concept and applied it, for instance, in the context of products in which quality is evaluated after purchase and experimentation, with asymmetric information between buyers and sellers (Gammoh et al., 2006). The studies from Kirmani, A. presented evidence that, when no other information is available, consumers may perceive the investment in advertising as a signal of quality (Kirmani, 1990). Klien et al.’s research assumed that consumers have access to many different digital information sources and introduced a new metric to evaluate cross-media exposure. This is important to understand consumers’ decisions in pre-purchase phases (Klein et al., 2020). Future research might further explore the diversification of investments in paid, owned and earned media, in alternative to massive paid media exposure.

4.9.1 Example of questions for future research.

- How to design a comprehensive consumer journey, that can include consumers’ decisions in pre-purchase phases and conciliate it with privacy regulation principles such as data minimization?
- How to reach consistent cross-media signals with the limitations for data processors to adopt identity resolution solutions?

4.10 Structuration theory

Quach et al. (2022) use the structuration theory as the conceptual body to extend marketing theory to data privacy context. The structuration theory embraces the interaction of structures (governmental agencies and regulators) and actors (firms, stakeholders) in social processes (Giddens, 1983). The studies from Quach et al. rely on the idea that data privacy...
Regulations enforce structure on consumers, firms and policymakers, generating interactions and responses by each key actor. Giddens’ structuration theory is useful in domains of organization science where it is important to study the interplay between actors, in the context of the structure where they belong, including social, ethical and physical aspects, and boundaries in the environments where the firm operates (Luo, 2006; Whitford and Zirpoli, 2014). Quach et al.’s integrated framework includes three tenets and seven propositions. For those researchers that would like to conduct their studies based on this contribution, there are some avenues to explore. First, it is hard to homogenize firm actions because the relationship with regulators and consumers might be very different and, in some cases, conflictual. In the case of advertising, there is a true supply chain where brands sit on the demand side, publishers and platforms on the supply side and data brokers in the middle, generating its own tension among them and different interplays with regulators and consumers. Second, the integrated framework from Quach et al. can be extended to study tensions from regulatory initiatives beyond privacy. For instance, regulatory initiatives to target the imbalance of market power between platforms and publishers have been studied by the academy with unexpected results for regulators, based on the evidence from Spanish and German cases (Calzada and Gil, 2020; Chiou and Tucker, 2017). This impacts the interplay of two different types of data providers (publishers and platforms) not only with the regulators but also with consumers in aspects so important as the free access to content and willingness from consumers to share their data in exchange.

4.10.1 Example of questions for future research.

- How can this integrated framework be extended to embrace the different roles of the firm?
- What will change in the dynamics of this structure if we incorporate other regulatory initiatives beyond data privacy?

4.11 Value theory for personal data

Spiekermann, S. and Korunovska, J. raised the question of whether personal data will still be available for marketers at near no cost in the future. These authors proposed a value theory for personal data from the perspective of how users value their personal data, grounded on constructs from economics, marketing, social and psychological ownership theories (Spiekermann and Korunovska, 2017). This research measured the consumers’ willingness to pay to protect their data and the willingness to accept money for data in an experiment with Facebook users. The study revealed a positive correlation between the level of engagement and value perceived for personal data when participants were confronted with a scenario where Facebook would be shut down and users could not access their information. Researchers interested in contributing to a value theory for personal data might be interested in exploring some considerations. For this experiment, the information on Facebook profiles included not only personal data but also content, such as pictures and photo albums. This is typical in the case of industry platforms exhibiting direct network effects. In the case of Facebook, this is visible by the way Facebook grabs friends of friends of users (Gawer, 2020). Platforms that constitute two-sided markets with asymmetric externalities tend to always incent the platform to collect and analyse personal data, although the degree of importance that consumers grant to personal data might impact the efficiency of data collection processes and commercial practice of platforms (Duan et al., 2022). However, there are different kinds of organizations that can be walled gardens of personal data. For instance, the publishers can also collect data valuable for marketers. And
all the programmatic advertising ecosystem in general only depends on taxonomists to create useful audiences for marketers with data usually collected based on cookies. In these cases, the constructs related to willingness to pay will not apply because there is no user profile, data or content that consumers would be willing to pay to preserve it. In addition, the sense of psychological ownership that would influence the willingness to accept money for data of those users with higher levels of engagement would not apply in these cases as well.

4.11.1 Example of questions for future research.

- Can value theory for personal data be applicable when there is no data or contents that consumers would be willing to pay to preserve it?
- What other metrics might be used to measure how users value their personal data?

5. Conclusion

DDM evolved over the past decades with no significative restrictions. Innovations and market autoregulation were the main responsible for the landscape we know today. But a new generation of regulations is changing the landscape dramatically. Most came to force less than five years ago, raising impact and academic research all over the world as described in Figure 3. Data privacy regulations are being complemented by data sovereignty rules and anti-trust and competition regulations. For marketing researchers, the topics related to DDM needs more attention, as the application of data in marketing fields is far more studied in the fields of computer science and computer-mediated interactions (Bucklin and Sismeiro, 2009). For firms, it is important to raise data governance at the level of strategic management. Organizations must be capable and savvy to explore the full value of data, being able to measure it and reduce the associated risks. The implementation of business ethics principles, media strategy, data ownership, strategic partnerships, decisions on internalization versus outsourcing and the development of competencies and learning capabilities are areas of paramount importance for management in a regulated world.

References


About the authors
Jorge Xavier is a PhD candidate in management at ISEG, Lisbon School of Economics and Management, at the University of Lisbon. He holds a MSc in Science, Technology and Innovation from the University of Aveiro and published papers in the Information Society field. Xavier is an Invited Professor to IPAM Marketing School, on the postgraduate program in Marketing Research, Intelligence and Analytics. He has a professional background with 20 years of experience in digital marketing and advertising technologies, working for leading companies including Microsoft and Oracle and developing projects with brands and publishers. His current research interests also embrace data-driven organizations in the health-care industry. Jorge Xavier is the corresponding author and can be contacted at: jorge.xavier@phd.iseg.ulisboa.pt

Winnie Ng Picoto is an Associate Professor of information systems and operations management at ISEG, Lisbon School of Economics and Management, at the University of Lisbon. She holds a BA in industrial engineering and management from the Instituto Superior Tecnico, a MIS from ISEG and a PhD in management from the Technical University of Lisbon. She is a member of the Advance Research Centre. Her previous work experience includes information systems consulting. Her current research interests include the use of innovative IS, IT value, big data and emerging technologies. Her work has been published in journals such as *European Journal of Information Systems, Journal of Business Research, Industrial Management and Data Systems* and *Journal of Organizational Computing and Electronic Commerce*.

For instructions on how to order reprints of this article, please visit our website: [www.emeraldgrouppublishing.com/licensing/reprints.htm](http://www.emeraldgrouppublishing.com/licensing/reprints.htm)
Or contact us for further details: permissions@emeraldinsight.com