Unlocking value from AI in financial services: strategic and organizational tradeoffs vs. media narratives

Gianvito Lanzolla[1], Simone Santoni[1] and Christopher Tucci[2]
[1]The Business School (formerly Cass) – City, University of London, London, UK
[2]Imperial College Business School – Imperial College, London, UK

Abstract

One of the sectors that AI is expected to transform more radically is financial services. In this Chapter, first, building on the extant literature, we develop a general conceptual model of the strategic and organisational tradeoffs inherent in extracting value from AI. Second, we use a topic modelling approach on two sources of data (a "corpus" of articles published in the global business press and a "corpus" of web articles concerning AI in financial institutions located in the City of London) to identify the narratives about AI in financial services from 2013 to 2019. The comparison between our model of AI value creation and the narratives emerging from our topic models reveals a significant divide. We elaborate on the implications of this divide and we offer suggestions on how to manage AI adoption for value creation.

Keywords: AI; business model; strategic tradeoffs; organisational tradeoffs; AI media narratives; financial services

Introduction

In 1955, McCarthy wrote that "The goal of artificial intelligence is to develop machines that behave as if they were intelligent." If we consider the last 65 years, Artificial Intelligence (AI) is now on its third, or maybe fourth, wave of excitement and promises. The previous waves have ended in disappointment. In the current wave, algorithms powering AI have improved, and data and computing power are more readily available, and at a relatively cheaper cost/unit. Is it the right time for AI to deliver on its promises of value creation? Or are we heading towards another "winter" of AI?

One of the sectors that AI is expected to transform, or even disrupt, more radically is the financial services sector. For instance, Hawksworth et al. (2018) indicate that the jobs most vulnerable in the earliest stages of AI automation (the "algorithmic wave") are in financial services. According to Stefanel and Goyal (2019), financial services account for 19% of total IT spending on artificial intelligence, and AI is expected to generate US$1.2 trillion of additional value for the financial industry by 2035. Walch (2020) reports that in 2018 alone, over $9.3 billion were raised by AI startups. However, anecdotal and emergent empirical evidence show that leveraging AI for value creation has proven harder than expected. According to several studies, executives find it hard to assess the real business impact of AI, and even harder to scale it in their business models (e.g., McKinsey, 2018).

In this Chapter, we set out to shed some light on the mechanisms through which AI can transform / disrupt value creation in the financial sector, and on why extracting value is proving harder than expected. Thus, **first,** we develop a **conceptual model of AI value creation through the lens of the business model construct.** The business model is a construct—and a narrative—that focuses on the interdependencies among firm activities that lead to value creation (Massa et al., 2017; Lanzolla and

Markides, 2020). AI is expected to influence value-creating activities in several ways (Brynjolfsson & Macafee 2017). AI is expected to create value by replacing and complementing many activities that are performed by financial institutions, e.g., AI-powered approval of loan applications that replaces the work of loan managers. Furthermore, AI is also expected to enable the creation of brand-new ways of delivering on financial services' legacy activities. Consider, for instance, robo-advisory or peer-to-peer lending. It follows from these considerations that for its focus on interdependent activities and value creation, the business model construct is a suitable platform to systematically analyze the impact of AI on value creation. We present our model in Section 2.

Second, we undertake a study of how **artificial intelligence** has been associated with transformation / disruption in the **financial services sector** over the 2013-2019 time period. To deliver on this goal, we use a topic modeling approach on two sources of data: a "corpus" of articles published in the global business press (in the English language) concerning AI in financial services; and a corpus of articles published on web pages concerning AI in financial institutions located in the City of London, one of the major financial centers in the world. The topic modeling on the global business press corpus of data reveals general (global) narratives about AI in financial services. The second study on the Internet corpus of data reveals the narratives related to financial institutions located in the City of London. Our topic models document that in the last ten years, the narratives around AI have been mainly focused on technology, talent, and technology venturing. We present our topic models in Section 3.

Third, in the last section of this Chapter, **we compare and contrast our model of AI's value creation with the narratives emerging from our topic models**. Our analyses reveal that there is a significant divide between the strategic and organizational tradeoffs that should be made to use AI for value creation and the narratives around AI in the media. Media narratives have a significant role in influencing management attention, and therefore the development of business cases and resource allocation (e.g., Rindova et al., 2006; Tripsas, 2009; Lanzolla and Suarez, 2012). We conclude that perhaps one of the reasons for AI still being a long way from delivering on its potential is the chasm between the tough choices necessary to leverage AI for value creation and these somewhat generic—or even overly simplistic—narratives. We offer suggestions on how to close this divide by commenting on AI adoption timing and market entry timing.

Towards a business model-oriented classification of AI's value creation potential

Business models

Over the last two decades, the business model has become an increasingly important concept, particularly in the fields of technology and innovation management (Massa & Tucci, 2014; Tripsas & Gavetti, 2000), strategy (e.g., Casadesus-Masanell & Zhu, 2013; Teece, 2010; Bigelow and Barney, 2020; Lanzolla and Markides, 2020) and, more recently, environmental sustainability (Schaltegger, Lüdeke-Freund, & Hansen, 2012) and social entrepreneurship (e.g., Seelos & Mair, 2007). At a very general and intuitive level, *a business model is a description of an organization and how that organization functions in achieving its goals (e.g., profitability, growth, social impact, …)* (Massa et al., 2017). In the management literature, definitions of a business model have then ranged from "stories that explain how enterprises work" (Magretta, 2002, p.4) to "a system of interdependent activities that transcends the focal firm and spans its boundaries" (Zott & Amit, 2010, p. 216) to "the logic, the data and other evidence that support a value proposition for the customer, and a viable structure of revenues and costs for the enterprise delivering that value" (Teece, 2010, p. 179).

Business model narratives can be descriptive, but they might also take on a normative role as well; in other words, provide aspirational vision for how an organization should work rather than how it does work. Narratives themselves can be thought of as an important tool to provide meaning from ambiguous situations and persuade audiences that might be skeptical that the account of reality developed by the narrator is credible (Garud & Giuliani, 2013). Perkmann & Spicer (2010) propose that due to their forward-looking nature, business model narratives may be an important factor in shaping the expectations of different stakeholders about how things might evolve in an organization's future. Executives and entrepreneurs can construct business model narratives and utilize them to both reduce cognitive complexity as well as to achieve certain goals. As a communications device, a business model narrative may be useful in influencing external stakeholders; creating legitimacy for a new business idea (for a startup or a corporate venture) for example by demonstrating similarities between the new business model and an acknowledged successful business model; or even "nudging" actors toward different social actions, such as by making certain decision criteria more salient or helping agents decide how to act (Massa and Tucci, 2014).

Focusing on business models as a system of interdependent activities, Lanzolla and Markides (2020) propose that the consideration of such interdependencies provides a roadmap to: (1) develop new insights on how to build superior strategies; and (2) explain company performance variation especially when heterogeneity in resources and capabilities is not strong and barriers to imitation are weak.

For its focus on the narratives and the web of interdependent activities which lead to value creation and value capturing, firms' business models are a level of analysis ideally suited to understanding the impact of nascent technologies such as Artificial Intelligence.

Towards a business model-oriented classification of AI's value creation potential

Artificial Intelligence does much more than automate and "informate" (Bailey et al, 2019). Ongoing developments are leading to the emergence of intelligent technologies that somehow mimic or even outperform humans in a wide variety of skilled and cognitive acts. For example, intelligent technological actors are increasingly performing work such as collecting and processing information; dividing, assigning, and integrating tasks; allocating resources; and making decisions (e.g., Faraj et al. 2018). Agrawal et al. (2018) point out that AI is a collection of technologies that causes the cost of prediction to drop, expanding the opportunities to use prediction models while at the same time growing the importance of data, judgment, action, and data-driven decisions.

There are different expected benefits from AI adoption, which we synthesize into an articulation of different levels of expectations. We use these different levels to build up the Y-axis in Figure 1 below, which we label "AI's intended benefits." At the origin, the lowest level is no benefit whatsoever. At a certain level of adoption, digitalization and AI should be expected to accomplish the bare minimum of predicting and automating simple business processes, thus automating standalone decisions for simple tasks and increasing the efficiency of those decisions. As we move up the Y-axis, the next level is to apply AI for automating predictions and decision making on more complex tasks, i.e., tasks where there are interdependencies or coordination needs across tasks. Finally, at a high level of adoption and use, the benefits could be characterized by continuous learning, in which feedback and learning is incorporated into a "learning loop." Thus, we summarize the impact of AI on or benefits to organizations as efficiency in decision making, coordination processes, and organizational learning as shown on the Y-axis of Figure 1.

** Place Fig. 1 about here **

Irrespective of the expected benefits represented on the Y-axis, we can also think of a "maturity model" of AI adoption (e.g., Alsheibani et al., 2018) in which AI is increasingly adopted within an organization and has a broader and broader impact on the firm's business model. In some sense, this corresponds with the idea of increasing business model innovation in different components of a business model (Foss and Saebi, 2017; Markides, 2013). We label this "AI's impact on the firm's activity system,"—consistent with the idea that business models describe a system of activities both within an organization and between an organization and its external ecosystem (Massa et al., 2017; Casadesus-Masanell and Ricart, 2012; Zott and Amit, 2010)—and we represent it on the X-axis of Figure 1. The levels of this dimension start from no impact whatsoever near the origin. The next level that we articulate is that AI could be used to complement, or augment existing activities, for example by providing suggested predictions or support (Liebowitz., 2020). Once that has been achieved, AI is used to replace some current activity or activities as part of the automation process mentioned above, similar to IT, Internet, and Industry 4.0 automation processes (Brews and Tucci, 2003). Finally, at higher levels of integration, AI could be used to enable or create entirely new activity systems; in other words, reconfigure or innovate the business model (cf. Massa and Tucci, 2014; Markides, 2013). It follows that in terms of the scope of the business model change, AI can complement existing activities, replace existing activities, and enable the delivery of brand-new activities, as represented on the X-axis.

In the pursuit of value extraction from AI, Fintech, Big Tech, and legacy financial institutions are all grappling with the two dimensions described in Figure 1. For example, Fintechs are often focused on the X-axis and in deploying AI-based business models which (might) make legacy business models obsolete; whereas legacy banks or insurance companies are hoping to complement or replace legacy activities, with a focus on efficiency.

Making it happen: tradeoffs for value creation

As is often the case for any new technology, the adoption of AI requires making tradeoffs. In other words, firms will not automatically extract value from AI just by adopting it. In this section, building on the extant management literature, we elaborate four idiosyncratic tradeoffs that companies should be making and managing when adopting AI.

Tradeoff 1: efficiency vs. slack. AI poses the tradeoff between (increasing) efficiency vs. (decreasing) organizational slack. On one hand, as hinted above, AI equips organizations with capabilities such as monitoring, control, and addressability that increase the scope for optimization and efficiency (cf. Brews & Tucci, 2004; 2007). However, on the other hand, through AI, different tasks' processes and functions become interconnected and inseparable from one another (Luo et al., 2012) and this might trigger—perhaps unintentionally—consequences such as decreasing diversity and organizational slack (e.g., Lazer & Friedman, 2007). Thus, we see the fundamental tradeoff between efficiency and organizational slack. Companies build efficiency, and this may give them the impression of becoming more competitive. Yet, this happens at the expense of slack, which makes organizations inherently less innovative and resilient to disruptions or shocks (Lanzolla, Pesce and Tucci, 2020).

Tradeoff 2: coordination vs surveillance. Another tradeoff is evident between facilitating coordination and facilitating surveillance based on AI technologies. Artificial intelligence and digitalization in general can certainly make it easier to coordinate task accomplishment across larger and larger groups. A simple example starts with the idea of micro-tasks in crowdsourcing, e.g., Amazon Mechanical Turk, where participants or employees do a small part of the work, and then algorithms aggregate the information and (re)assemble it into a whole (Bernstein, 2012; Little, Chilton, Goldman and Miller, 2010). At a more complicated level, some types of human coordination can be automated partially or even entirely: think about software that automatically schedules meetings, or back-office processes that query databases and integrate information from multiple sources (Davis, 2016).

Puranam (2018) highlights the economic efficiency benefits to improved digitally mediated coordination.

However, as Kellogg et al. (2020) note in their extensive review, these kinds of aids to coordination come at a cost, and a good part of that cost is represented by increased surveillance along several dimensions. In order to partially or fully automate coordination, the “system” has to collect data from employees. Referring back to the simple example of automatic meeting scheduling, the employer would need everyone to use the same calendar (diary) system and the calendar data would need to be visible centrally. The natural next step might be for employers to compare how many meetings each employee has per week, or which employees are spending their time “profitably,” or even publicize the meeting productivity of each employee so they can see each other’s statistics. These are examples of different kinds of control but mainly refer to technical control, where new technologies are used instead of direct supervision to monitor and control employees (Braverman, 1974; Aiello and Svec, 1993). This leads to feelings of constantly being under surveillance, which may lead to alienation, resistance, and even sabotage (Kellogg et al., 2020). Thus, we see the fundamental tradeoff between coordination and surveillance. Companies need some kind of surveillance to improve or automate coordination using AI and digital technologies. Whether they then use it as a means of control is an open question, but many scholars would agree that once you start collecting data, it is difficult to resist the temptation (Davis, 2016; Schafheitle et al., 2020).

Tradeoff 3: learning vs. propagation of “black boxed” and “biased” learning. As highlighted in the previous paragraphs, AI has opened up new ways of carrying out “old” activities and created a set of technological solutions in search of “new” activities (Cockburn, Henderson & Stern, 2018). Concerning new ways of carrying out old activities, there is abundant work that documents how AI has been applied in financial services for organizational learning. For example, Natural Language Processing has become a popular tool to extract information from contracts, such as swap expirations, and notify clients promptly (Mik, 2017); Machine Learning has unlocked the value of portfolio optimization by attenuating some long-standing estimation issues (Ban, Karoui & Lim, 2018); Deep Learning has turned satellite images of night-time lights into key features to predict economic and social trends at a “local” scale (Proville, Zavala-Araisa & Wagner, 2017).

Regardless of the specific scope of adoption, a possible way to appreciate the effects of AI, we posit, is to consider the complex, multi-level interaction between a firm's experience and the context where the experience takes place (Argote & Miron-Spektor, 2011), namely, the social and organizational situations in which AI algorithms are introduced and used. Such an analytical framework allows us to highlight the role of two crucial factors that could shrink the payoff to AI, that is, “blackboxing”[*] (Latour 1999) and “algorithmic bias” (Lambrecht & Tucker, 2019).

Some Machine Learning and, especially, Deep Learning algorithms have been labeled as “black-box” machines due to their inherent complexity (Rudin, 2019). The lack of interpretability (Shrestha, Ben-Menahem & von Krogh, 2019) is another element that accounts for the existence of black-box machines. Organizational theorists emphasize the adoption of AI technologies as a process that is situated in the context of concrete occupations (e.g., Bailey & Leonardi, 2015). Most workers do not fully grasp what kinds of data are being collected about them, how they are being used, or how to contest them (see also Bolin et al., 2015). These elements reinforce the idea that the adoption of AI presents both learning opportunities and black-boxing risks, which are augmented by the lack of familiarity with / trust in algorithms on the part of professionals.

[*] Latour defines black-boxing as *"the way scientific and technical work is made invisible by its own success. When a machine runs efficiently, when a matter of fact is settled, one needs to focus only on its inputs and outputs and not on its internal complexity. Thus, paradoxically, the more science and technology succeed, the more opaque and obscure they become."*

A second factor that could limit the impact of AI is algorithmic bias, i.e., the fact that trained models reflect the values, beliefs, and norms of their creators (Crawford, 2016). For example, facial recognition software embedded in most smartphones works best for those who are white and male (Buolamwini & Gebru. 2018); Deep Learning applications for computer-aided diagnosis can show heterogeneous performance across female and male patients (Larrazabal et al. 2020); Amazon, whose global workforce is 60 percent male and where men hold 74 percent of the company's managerial positions, discontinued use of a recruiting algorithm after discovering gender bias (Vincent, 2018).

Tradeoff 4: employee empowerment vs. employee disengagement. On one hand, AI might allow a reduction in cognitive load by minimizing the effects of stress and time pressure, while bringing the multitude of variables outside employees' control under management (Bailey, Leonardi, and Barley 2012). On the other hand, there is significant evidence that operating in the physical world through digital interfaces prompts changes in the organization of work, alters the way people make sense of—and come to trust—the objects with which they work, and can increase cognitive load. In this vein, Zuboff (1988) shows that in paper mills, before digitalization, workers relied on their senses to get information about the production process. In Zuboff's seminal study of paper mills (1988), she proposes that the implementation of the new control system triggered emotions such as anger and fear in some workers, while others documented the presence of emotions such as happiness, joy, relief, frustration, irritation, and annoyance. Rafaeli and Vilnai-Yavetz (2004) point out that digital artifacts may trigger emotional reactions from individuals when they interrupt routines. Specifically, emotions play a large role in the period between the moment the routine is interrupted, and the time new routines are established (or old routines are reestablished). Lerner and Keltner (2000) explore the effects of emotions occurring prior to the deployment of new digital technology. They suggest that emotions are triggered based on users' expectations of how the new technology will affect them, their work, their performance, and their coping mechanisms.

Narratives about AI in financial services business models

In this section, we use a topic modeling approach to reveal how AI has been associated with value creation in financial services[†] by using two sources of data: articles published in the global business press (in the English language) concerning AI in financial services; and articles published in web pages concerning AI emanating from financial institutions located in the City of London, one of the world's major financial centers. In particular, we focused on the 100 largest companies[‡] in terms of revenues (Appendix B reports the full list of companies included in the sample). In this second study, we explore how a general audience represents the positioning of individual companies in relation to digital technologies. We also explore how the positioning of each individual firm has changed over time. Below we describe in detail our data sources, methods, and the findings of the two studies.

Data sources and methods

Business Newspapers Corpus. Figure 2 illustrates the inter-temporal distribution of roughly 5000 newspaper articles. All articles deal with some facets of the "artificial intelligence" phenomenon and focus on the context of financial services (Appendix A provides an account of the data gathering process, including the set of keywords and other criteria we used to sample relevant articles). Between January 2000 and December 2012, both *The Financial Times* and *The Wall Street Journal*—arguably two of the most prominent business newspapers—paid very limited attention to the role of digital technologies in financial services. For example, in 2012 we were able to find only 65 articles on the

[†] We considered SIC codes 64 "Financial service activities, except insurance and pension funding" and 65 "Insurance, reinsurance and pension funding, except compulsory social security."

[‡] For the purpose of data gathering, separate legal entities associated with the same company (e.g., "Lloyds Bank Asset Finance Limited" and "Lloyds Bank PLC") were grouped within the same entity and treated as a single company (e.g., "Lloyds").

subject. From January 2013 onwards, the two newspapers started to cover the topic of "digital technologies & financial services" on a more systematic basis. In a period of four years (January 2013 - December 2016), the number of articles devoted to the topic rose by more than 300%. Economic journalists' attention to the topic reached its maximum in 2017 when over one thousand articles were published.

** Place Fig. 2 about here **

Internet Corpus about City of London firms. In a second stage, we conducted a broad crawl and fetched the contents of the top 100 search results that Google returned for each company-year-keyword combination. Then, we used topic modeling to analyze the resulting corpus of text, which occupies circa 21.7 GB of storage space. Figure 3 visually depicts the inter-temporal distribution of web pages retained for the analysis (Appendix B describes some key aspects of the research design, including the keywords passed to Google Search, the criteria adopted to sample web pages, and the Natural Language Processing that lies behind the topic modeling).

** Place Fig. 3 about here **

Methods. To uncover the themes discussed in this corpus of economic newspaper articles, we decided to rely on topic modeling (Blei et al. 2003), an unsupervised machine learning application that identifies clusters of terms that tend to co-occur in the same document (readers who would like a more detailed description of the estimation procedures we carried out can refer to the "Data Analysis" section of Appendix A). Relative to other unsupervised approaches to the analysis of meaning in natural language, topic modeling requires very few assumptions and limited knowledge of the corpus of text at hand. For example, semantic network analysis implies scholars have (theoretical) expectations about the specific meanings that are reflected in the data and the affiliation of words with meanings. Topic modeling facilitates the process of discovering the most relevant meanings in the corpus by revealing the existence of clusters of words that co-occur within and across units of text. Moreover, scholars can embrace a generative approach to data analysis by efficiently estimating and comparing alternative topic models, i.e., models with different numbers of unique clusters of words, or iterating over concrete examples and results to make sense out of the clusters of words included in the corpus of text.

Findings of study 1 – Narratives about AI associated with financial services globally

Table 1 reports the structure of hidden themes in the corpus. Each column of the table corresponds to a topic. Rows indicate the ten terms that are most likely associated with each individual topic. Terms are ranked in descending order of importance, i.e., strength of association. At the intersection of row i and column j, we report the i^{th} most important word for topic j (e.g., "machine") along with the strength of the association between the word and the topic (e.g., 0.32).

** Place Table 1 about here **

Topic #1, reported on the far-left hand side of the table, concerns the role of AI in personal finance; Topic #2 emphasizes the market expectations of industry experts *vis-à-vis* the diffusion of AI in the financial services sector; Topic #3 deals with regulatory aspects in banking; Topic #4 highlights the linkages between AI and customers in the insurance sector; Topic #5 focuses on the strategic role of AI in the banking sector; Topic #6 concerns the investment management sector; Topic #7 represents the financing of new, tech-based ventures; finally, Topic #8 relates to the private equity field. We note here that while it is possible to identify a relatively large group of topics that reflect the general business context wherein AI is deployed (see Topics # 1, 2, 3, 5, 6 and 8) only two topics seem to be directly related to AI adoption – i.e., topic #4 (customer analytics) and topic #7 (technology ventures).

Figure 4 expands on the results reported in Table 1 to show the saliency[§] of each topic over the timespan of our analysis. The slope chart indicates the existence of a group of topics whose popularity decreases over time. Topic #2, concerning the relationship between digital technologies and, presumably, their impact on market expectations, is circa 30% more likely to appear in an article published in 2013 than in an article published in 2019. Three other topics (i.e., #1 – "personal finance"; #3 – "regulatory aspects in banking"; #5 – "strategic role of digital technologies in banking") seem increasingly less central over time. The analysis of the slope charts suggests also the existence of topics with positive trends. Particularly, Topics #4 – "digital technologies and insurance market" – and #7 – "technology ventures" – become increasingly more central over time; in 2019, they appear as "dominant topics" in one third of all documents.

** Place Fig. 4 about here **

Findings of study 2 – Narratives about AI associated with City of London financial service firms

Table 2 reports the results that emerge from the topic modeling in the form of a term-to-topic matrix (regarding the organization of this table, see the previous paragraph of this chapter). Topic #1 deals with the human capital aspect that underlies the adoption of digital technologies; Topic #2 concerns financial markets; Topic #3 highlights the coverage, on the part of different social and communication media, of digital-technology related activities; Topic #4 emphasizes the intersection of AI / digital technologies and various business and technological domains; Topic #5 concerns the business of banking and insurance; Topic #6 highlights the theme of machine learning and its methodological and technical bases; Topic #7 highlights the linkages between data/analytics and customers; finally, Topic #8 relates to cloud computing technologies. Overall, similar to the Study 1, the topics emerging from this analysis seems to relate to the general business context and AI adoption related topics. However, different from Study 1, the general business context (Topics # 2, 3, and 5) are smaller in number to the AI adoption related topics (Topics # 1, 4, 6, 7 and 8).

** Place Table 2 about here **

[§] By saliency, we mean the probability to see the focal topic represented in a document randomly drawn from the corpus.

Figure 5 visually depicts the saliency[**] of the individual topics included in Table 2 over the timespan of our analysis. Topic #7, concerning the theme of digital technologies and value creation, stands out from the rest of topics—it is the most salient topic throughout the period of the timespan and its popularity, relative to other topics, increases over time. In 2019, Topics #2 (on AI and financial markets) and #5 (on the business side of banking and insurance) are amongst the most salient ones as a result of a positive-consistent trend over the entire time span. A group of three topics, namely, Topics #1 (focusing on the human capital side of AI), #3 (media coverage of initiatives relating to AI), and #6 (on machine learning) decline substantially over time. Topics #4 (highlighting the intersections of AI with specific business and technological domains) and #8 ("cloud computing") appear in a small yet stable group of documents.

** Place Fig. 5 about here **

Figure 6 elaborates the results of the topic modeling reported in Table 2 by mapping Topics #1 - #8 onto the companies included in the sample. In particular, we tag each company-year observation in terms of a dominant topic, i.e., the topic with highest probability to appear in a document covering company *i* at time *j*. This opens up the possibility of representing—and investigating—the changes that characterize a company's semantic positioning (with regard specifically to digital technologies). The figure constitutes of a series of 69 company-specific trajectories;[††] each company-year combination is color-coded according to the dominant topic.

** Place Fig. 6 about here **

In conclusion, our topic models demonstrate that in the last ten years, both in the business press corpus and in the Internet corpus centered on City of London firms, the narratives around AI clustered around two macro areas: general business context and (more) specific AI-related topics. The general business context topics seem to describe the broader business and regulatory implications related to AI diffusion. The (more) specifically AI-adoption-related topics focus on the description of the AI and digital technologies themselves (Topic #4 in Table 1 and Topics # 6 and 8 in Table 2); the modalities and challenges to adopt AI (Topic #7 in Table 1 and Topic #1 in Table 2); and the scope of AI adoption (Topic #4 in Table 1 and Topics #4 and 7 in Table 2).

Unlocking value from AI in financial services: taking stock

The comparative analyses of the tradeoffs for value creation discussed in Section 2 vs. the findings of our topic model analyses (Section 3) reveal some patterns. While the management literature has significantly advanced our understanding of the tradeoffs underpinning value creation through AI (these tradeoffs are summarized in the body of Figure 1), the media narratives around AI seem to focus on more generic or general topics. Media narratives have an important role in creating adoption

[**] By saliency, again we mean the probability to see the focal topic represented in a document randomly drawn from the corpus.

[††] Some companies do not present enough documents to estimate company-specific document-to-topic probabilities. These companies were deleted list-wise. This left us with a set of 69 companies.

bandwagons (e.g., Lanzolla and Suarez, 2010) and in influencing management attention structures and subsequent resource allocation (March & Olsen, 1976; Rindova et al., 2006; Ocasio, 1997).

Our analyses here reveal a disconnect between the tradeoffs that should be made to leverage AI for value creation and new business models, and (some of) the narratives influencing AI adoption. This disconnect can be linked to two potential outcomes in AI adoption patterns. First, this disconnect might influence the AI adoption business cases. For instance, media narrative might suggest that technology features, talent, and technology venturing (please refer to our Tables I and II) are the only success factors while our analyses suggest that companies should be making several other strategic and organizational tradeoffs, as summarized in Figure 1. It follows that media narratives might play a significant role in the development of incomplete AI adoption business cases.

Second, our analyses in Section 3 show that media narratives change over time. Lanzolla and Suarez (2012) note that technology users are aware that a new technology does not remain unchanged once it appears in the market, nor is it always used in the same way, and that this makes users skeptical in embarking on learning about the new technology if they are not convinced that it is worth the learning investment. Lanzolla and Suarez show further that the disconnect between technology adoption and technology use is higher when adoption is triggered by media hype since potential users tend to discount media-based adoption bandwagons, and place greater value on *user bandwagons*. This is because in user bandwagons: (a) the information relates to users, not adopters, of the technology, and it is therefore considered more relevant and reliable by prospective users; and (b) the information is "new," that is, relates to users realized during the time period of a firm's technology adoption decision. The changing media narratives over time documented in this study (please refer to Figures 4 and 5) seem to reflect the situation described above very closely, thus suggesting a divide between AI adoption and AI actual use.

Jointly, our findings here describe a scenario in which the disconnect between strategic and organizational tradeoffs and media narratives might trigger yet another wave of disappointing AI adoption and yet another "winter of AI." Yet, on the other hand, this time around, companies might be better equipped with the knowledge and practice to exploit the full potential of AI.

Implications for unlocking value from AI

First, it follows from our arguments above that to drive value, top management should not only start by building complete AI business cases (and our model summarized in Figure 1 might help envision more complete AI adoption cases) but also be aware of the need for aligning AI user expectations with the AI adoption case, and media narratives do not help deliver on this.

The second implication of our study is about market entry timing with AI-based new products and/or services. AI technology is changing very quickly while market adoption of AI solutions is still relatively smooth, a situation that Suarez and Lanzolla (2007) call "abrupt pace of technological change and smooth pace of market adoption." Consider for instance the very slow uptake of peer-to-peer lending or robo-advisory services. In contrast, consider the rapid development in data availability and computing power that fuels AI. Under these contingencies, the "isolating mechanisms" (cf. Lieberman and Montgomery, 1988) that underpin first mover advantages do not seem to be easily activatable (Suarez and Lanzolla, 2007). Suarez and Lanzolla show further that a rapid pace of technology evolution might favor later entrants by enabling leap-frogging of the first-mover advantage while a smooth pace of market evolution might enable first mover to build strong client switching costs. As such, an implication might be to enter the market with sufficiently developed solutions that allow the firm to set up switching cost by creating new product "categories" and/or by providing clients with products and services which have more favorable benefit / price ratios. However, to sustain this advantage, such a firm should consider that any advantage on the technology side may be leapfrogged and that it would be critical to keep up with the pace of technology development by injecting significant financial and managerial resources.

The third, broader, implication is about the role that AI should have in any future business model. There are many aspects of business model innovation for legacy companies: AI could contribute to the rewiring of business models by complementing and replacing legacy activities and/or by enabling the delivery of new activities. This might increase the efficiency and perhaps profitability of adopters, as discussed above, but would imply perhaps only marginal changes in the business model (cf. Foss and Saebi, 2019). This activity rewiring does create new value, but one may question the level and extent of innovativeness of the business model reconfigurations that are realized (cf. Massa and Tucci, 2014). The constant influx of venture capital in born-AI-startups or the use of AI for radically re-inventing business models such as the ones for credit scoring (e.g., Pay Pal and Ant Financial) may point towards a reality in which AI might not only make legacy business models more efficient, but also completely re-invent them by radically changing the benefit / price ratio. In other words, barring the current dissatisfactory results of some new digital business models—such as peer-to-peer lending and robo-advisory—the scope for using AI for re-inventing value in financial services may be still untapped, perhaps not even fully imagined.

Which brings us to an important question about the role of AI in financial services: should financial institutions rewire their business models to become essentially technology companies? Or should they learn how to use AI while creating value through subject-matter specific mechanisms? Incumbents in the financial service sector have been pouring billions into AI and digital technologies to build a digital core, e.g., the Development Bank of Singapore and virtually all the leading banking groups. Yet, Monzo, one of the UK unicorns (founded in 2015) that has based its valuation success on the use of AI in its business model, has in 2020 reverted to call itself a "retail bank" and started hiring more and more retail banking subject matter experts as opposed to AI experts. The jury is still out as to whether value will be created and captured by incumbent institutions, AI startups, or the likes of Google, Apple and Amazon, often referred to as Big Tech. We believe that in the AI-transformed world, these worlds will co-exist in integrated ecosystems where subject matter expertise of legacy financial institutions and AI-powered business models will complement each other. The ultimate winners in terms of (more) value creation and (more) value capture will depend on several factors, including how effectively and efficiently firms grapple with the strategic and organizational tradeoffs described in Figure 1.

References

Abadie, A., & Gardeazabal, J. (2003). The economic costs of conflict: A case study of the Basque country. *American Economic Review*, 93(1), 113–132. https://doi.org/10.1257/000282803321455188

Acemoglu, D., Laibson, D., & List, J. A. (2014). Equalizing superstars: The Internet and the democratization of education. *American Economic Review*, 104(5), 523-27.

Agrawal, A., Gans, J., & Goldfarb, A. (2018). *Prediction Machines: The Simple Economics of Artificial Intelligence*. Harvard Business Press.

Aiello, John R., and Carol M. Svec. (1993). "Computer monitoring of work performance: Extending the social facilitation framework to electronic presence." *Journal of Applied Social Psychology* 23.7: 537-548.

Alsheibani, S., Cheung, Y. and Messom, C., (2018). Artificial intelligence adoption: AI-readiness at firm-level. *Artificial Intelligence*, *6*, pp.26-2018.

Argote, L., & Miron-Spektor, E. (2011). Organizational learning: From experience to knowledge. *Organization Science*, *22*(5), 1123-1137.

Bailey, D. E., P. M. Leonardi, and S. R. Barley. (2012). The lure of the virtual. *Organization Science* 23 (5): 1485–504.

Ban, G. Y., El Karoui, N., & Lim, A. E. (2018). Machine learning and portfolio optimization. *Management Science*, 64(3), pp. 1136-1154.

Bernstein, E. S. (2012). The transparency paradox: A role for privacy in organizational learning and operational control. *Administrative Science Quarterly*, *57*(2), 181-216.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3(Jan), 993-1022.

Bigelow, L. S., & Barney, J. B. (2020). What can strategy learn from the business model approach?. *Journal of Management Studies*.

Braverman, H. (1974). *Labor and Monopoly Capital*. New York: Monthly Review.

Brews, P. J., & Tucci, C. L. (2003). Internetworking: building internet-generation companies. *Academy of Management Executive*, *17*(4), 8-22.

Brews, P. J., & Tucci, C. L. (2004). Exploring the structural effects of internetworking. *Strategic Management Journal*, *25*(5), 429-451.

Brews, P. J., & Tucci, C. L. (2007). The structural and performance effects of internetworking. *Long Range Planning*, *40*(2), 223-243.

Brynjolfsson, E. and Mcafee, A. (2017). The business of artificial intelligence. *Harvard Business Review*, pp.1-20.

Buolamwini, J., & Gebru, T. (2018, January). Gender shades: Intersectional accuracy disparities in commercial gender classification. In Conference on fairness, accountability and transparency (pp. 77-91).

Casadesus-Masanell, R., & Ricart, J. E. (2010). From strategy to business models and onto tactics. *Long Range Planning*, 43(2-3): 195-215.

Casadesus-Masanell, R., & Zhu, F. (2013). Business model innovation and competitive imitation: The case of sponsor-based business models. *Strategic Management Journal*, 34(4): 464-482.

Cockburn, I. M., Henderson, R., & Stern, S. (2018). The impact of artificial intelligence on innovation (No. w24449). National Bureau of Economic Research.

Crawford, K. (2016). Artificial intelligence's white guy problem. *The New York Times*, 25(06).

Davis, G. F. (2016). Can an economy survive without corporations? Technology and robust organizational alternatives. *Academy of Management Perspectives*, *30*(2), 129-140.

Faraj, S., Pachidi, S., & Sayegh, K. (2018). Working and organizing in the age of the learning algorithm. *Information and Organization*, *28*(1), 62-70.

Foss, N. J., & Saebi, T. (2017). Fifteen years of research on business model innovation: How far have we come, and where should we go? *Journal of Management*, *43*(1), 200-227.

Fuertes-Olivera, P. A. (2007). A corpus-based view of lexical gender in written Business English. *English for Specific Purposes*, 26(2), 219-234.

Garg, N., Schiebinger, L., Jurafsky, D., & Zou, J. (2018). Word embeddings quantify 100 years of gender and ethnic stereotypes. *Proceedings of the National Academy of Sciences*, 115(16), E3635-E3644.

Garud, R., & Giuliani, A. P. (2013). A narrative perspective on entrepreneurial opportunities. *Academy of Management Review*, *38*(1), 157-160.

Gregory, R. W., Henfridsson, O., Kaganer, E., & Kyriakou, H. (2020). The role of artificial intelligence and data network effects for creating user value. *Academy of Management Review*.

Gunsilius, F. F. (2020). Distributional synthetic controls. Unpublished working paper. 1–52.

Hawksworth, J., Berriman, R., & Goel, S. 2018. Will robots really steal our jobs? Report 180123-184242-RB-OS. PWC.

Lambrecht, A., & Tucker, C. (2019). Algorithmic bias? An empirical study of apparent gender-based discrimination in the display of STEM career ads. *Management Science*, 65(7), 2966-2981.

Little, G., Chilton, L. B., Goldman, M., & Miller, R. C. (2010, October). Turkit: Human computation algorithms on Mechanical Turk. In *Proceedings of the 23nd annual ACM symposium on User interface software and technology* (pp. 57-66).

Kellogg, K. C., Valentine, M. A., & Christin, A. (2020). Algorithms at work: The new contested terrain of control. *Academy of Management Annals*, *14*(1), 366-410.

Lanzolla, G., & Markides, C. (2020). A business model view of strategy. *Journal of Management Studies*, in press.

Lanzolla, G., Pesce, D., & Tucci, C. L. (2020). The digital transformation of search and recombination in the innovation function: Tensions and an integrative framework. *Journal of Product Innovation Management*.

Lanzolla, G., & Suarez, F. F. (2012). Closing the technology adoption–use divide: The role of contiguous user bandwagon. *Journal of Management*, *38*(3), 836-859.

Larrazabal, A. J., Nieto, N., Peterson, V., Milone, D. H., & Ferrante, E. (2020). Gender imbalance in medical imaging datasets produces biased classifiers for computer-aided diagnosis. *Proceedings of the National Academy of Sciences*, 201919012. https://doi.org/10.1073/pnas.1919012117

Latour, B. (1999). *Pandora's Hope: Essays on the Reality of Science Studies*. Cambridge: Harvard University Press.

Lazer, D., and A. Friedman. 2007. The network structure of exploration and exploitation. *Administrative Science Quarterly* 52 (4): 667–94.

Liebowitz, J. (Ed.). (2020). *Data Analytics and AI*. CRC Press.

Leonardi, P. M. (2012). *Car Crashes without Cars*. Cambridge: MIT Press.

Lerner, J. S., & Keltner, D. (2000). Beyond valence: Toward a model of emotion-specific influences on judgement and choice. *Cognition & Emotion*, *14*(4), 473-493.

Luo, Y., S. L. Wang, Q. Zheng, and V. Jayaraman. 2012. Task attributes and process integration in business process offshoring: A perspective of service providers from India and China. *Journal of International Business Studies* 43 (5): 498–524.

Magretta, J. (2002). Why business models matter. *Harvard Business Review*, May: 86-93.

March, J. G., & Olsen, J. P. (1976). *Ambiguity and Choice in Organizations*. Bergen, Norway: Universitetsforlaget.

Markides, C. C. (2013). Business model innovation: What can the ambidexterity literature teach us? *Academy of Management Perspectives*, *27*(4), 313-323.

Massa, L., & Tucci, C. L. (2014). Business model innovation. In *The Oxford handbook of innovation management*, 20(18): 420-441.

Massa, L., Tucci, C. L., & Afuah, A. (2017). A critical assessment of business model research. *Academy of Management Annals*, 11(1): pp. 73-104.

McCallum, A.K. (2002). *MALLET: A Machine Learning for Language Toolkit*. (http://mallet.cs.umass.edu)

McKinsey & Company (2019). Global AI survey: AI proves its worth, but few scale impact. (https://www.mckinsey.com/featured-insights/artificial-intelligence/global-ai-survey-ai-proves-its-worth-but-few-scale-impact)

Mik, E. (2017). Smart contracts: terminology, technical limitations and real world complexity. *Law, Innovation and Technology*, 9(2), 269-300.

Morris, M. W., & Moore, P. C. (2000). The lessons we (don't) learn: Counterfactual thinking and organizational accountability after a close call. *Administrative Science Quarterly*, 45(4), 737-765.

Ocasio, W. (1997). Towards an attention-based view of the firm. *Strategic Management Journal*, *18*(S1), 187-206.

Perkmann, M., & Spicer, A. (2010). *What are business models? Developing a theory of performative representations*. Emerald Group Publishing Limited.

Proville, J., Zavala-Araiza, D., & Wagner, G. (2017). Night-time lights: A global, long term look at links to socio-economic trends. *PloS One*, 12(3), e0174610.

Puranam, P. (2018). *The microstructure of organizations*. Oxford: Oxford University Press.

Rafaeli, A., & Vilnai-Yavetz, I. (2004). Emotion as a connection of physical artifacts and organizations. *Organization Science*, *15*(6), 671-686.

Rasekhschaffe, K. C., & Jones, R. C. (2019). Machine learning for stock selection. *Financial Analysts Journal*, 75(3), 70-88.

Řehůřek, R., & Sojka, P. (2010). Gensim — Software Framework for Topic Modelling with Large Corpora. *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks* (http://is.muni.cz/publication/884893/en).

Rindova, V. P., Pollock, T. G., & Hayward, M. L. (2006). Celebrity firms: The social construction of market popularity. *Academy of Management Review*, *31*(1), 50-71.

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(5), 206–215. https://doi.org/10.1038/s42256-019-0048-x

Schafheitle, S., Weibel, A., Ebert, I., Kasper, G. Schank, C., & Leicht-Deobald, U. (2020). No stone left unturned? Towards a framework for the impact of datafication technologies on organizational control. *Academy of Management Discoveries*.

Schaltegger, S., Lüdeke-Freund, F., & Hansen, E. G. (2012). Business cases for sustainability: the role of business model innovation for corporate sustainability. *International Journal of Innovation and Sustainable Development*, *6*(2), 95-119.

Seelos, C., & Mair, J. (2007). Profitable business models and market creation in the context of deep poverty: A strategic view. *Academy of Management Perspectives*, *21*(4), 49-63.

Shrestha, Y. R., Ben-Menahem, S. M., & von Krogh, G. (2019). Organizational Decision-Making Structures in the Age of Artificial Intelligence. *California Management Review*, 66–83. https://doi.org/10.1177/0008125619862257

Stefanel, M., & Goyal, U. (2019). *Artificial Intelligence & Financial Services: Cutting through the Noise*. Report, Apis Partners, London.

Suarez, F. F., & Lanzolla, G. (2007). The role of environmental dynamics in building a first mover advantage theory. *Academy of Management Review*, *32*(2), 377-392.

Teece, D. J. (2010). Business models, business strategy and innovation. *Long Range Planning*, *43*(2-3), 172-194.

Thomke, S. H. (2003). *Experimentation Matters: Unlocking the Potential of New Technologies for Innovation*. Boston: Harvard Business Press.

Tripsas, M. (2009). Technology, identity, and inertia through the lens of “The Digital Photography Company.” *Organization Science*, *20*(2), 441-460.

Tripsas, M., & Gavetti, G. (2000). Capabilities, cognition, and inertia: Evidence from digital imaging. *Strategic Management Journal*, *21*(10-11), 1147-1161.

Vincent, J. (2018). Amazon reportedly scraps internal AI recruiting tool that was biased against women. *The Verge*.

Walch, K. (2020). “Is venture capital investment in AI excessive?” *Forbes*, 05 January. https://www.forbes.com/sites/cognitiveworld/2020/01/05/is-venture-capital-investment-for-ai-companies-getting-out-of-control, viewed 29 September 2020.

Zott, C., & Amit, R. (2010). Business model design: an activity system perspective. *Long Range Planning*, *43*(2-3), 216-226.

Zuboff, S. (1988). *In the Age of the Smart Machine*. New York: Basic books.

Exhibits

Fig. 1 – Unlocking value from AI: strategic and organizational tradeoffs

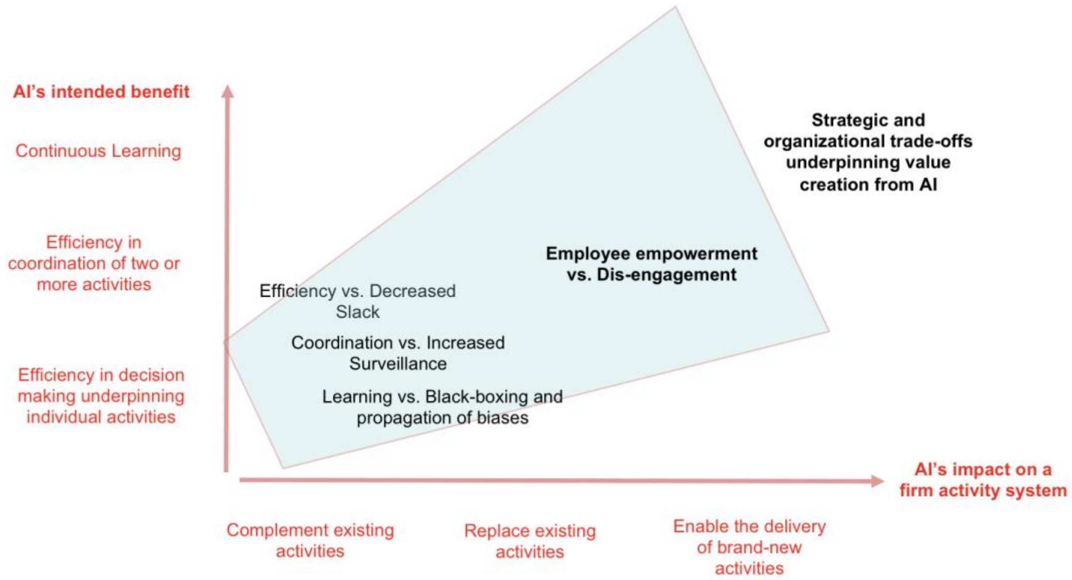
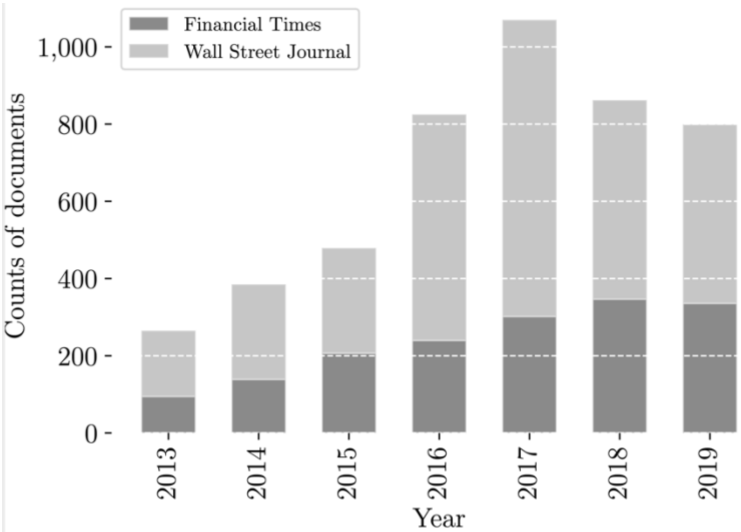
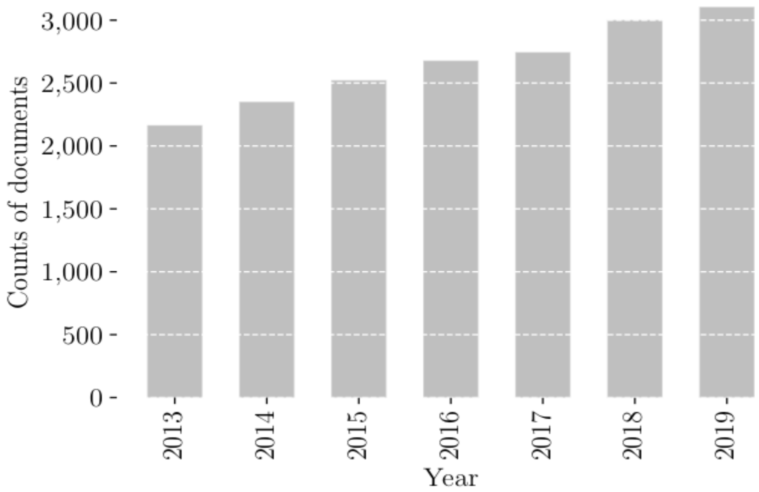


Fig. 2. Temporal distribution of articles included in the "economic newspaper corpus" (N = 5,036).‡‡



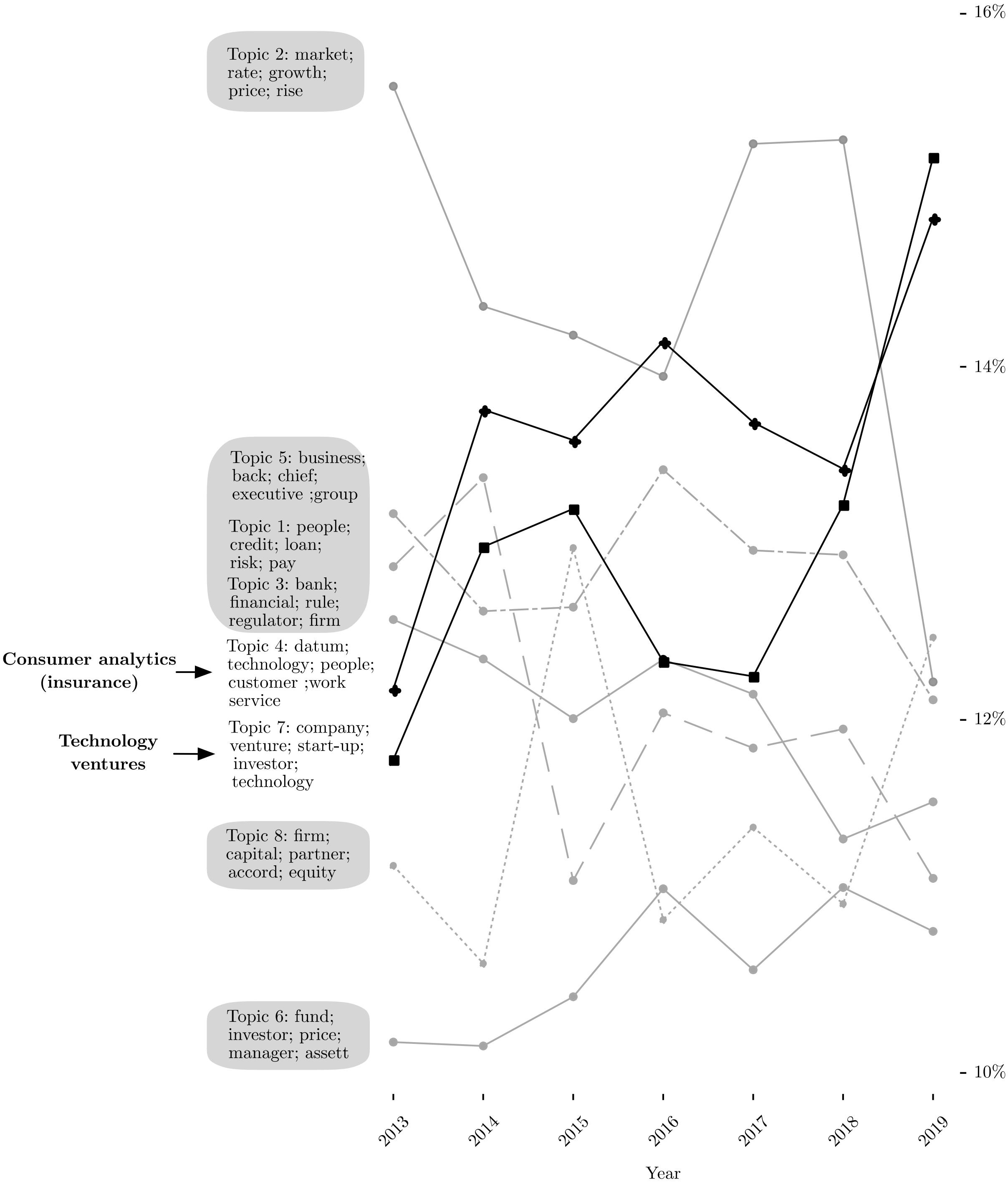
‡‡ The corpus concerns two economic newspapers, namely, the *Financial Times* and the *Wall Street Journal*. Articles were retrieved using the news aggregator Factiva on the basis of the following keywords: "artificial intelligence" or "deep learning" or "machine learning" or "big data" or "natural language processing" or "analytics." In order to be included in the sample, articles have to deal with the financial service sector (as per the categorization provided by Factiva).

Fig. 3. Temporal distribution of items for the “Internet” corpus (N = 21,612).[§§]



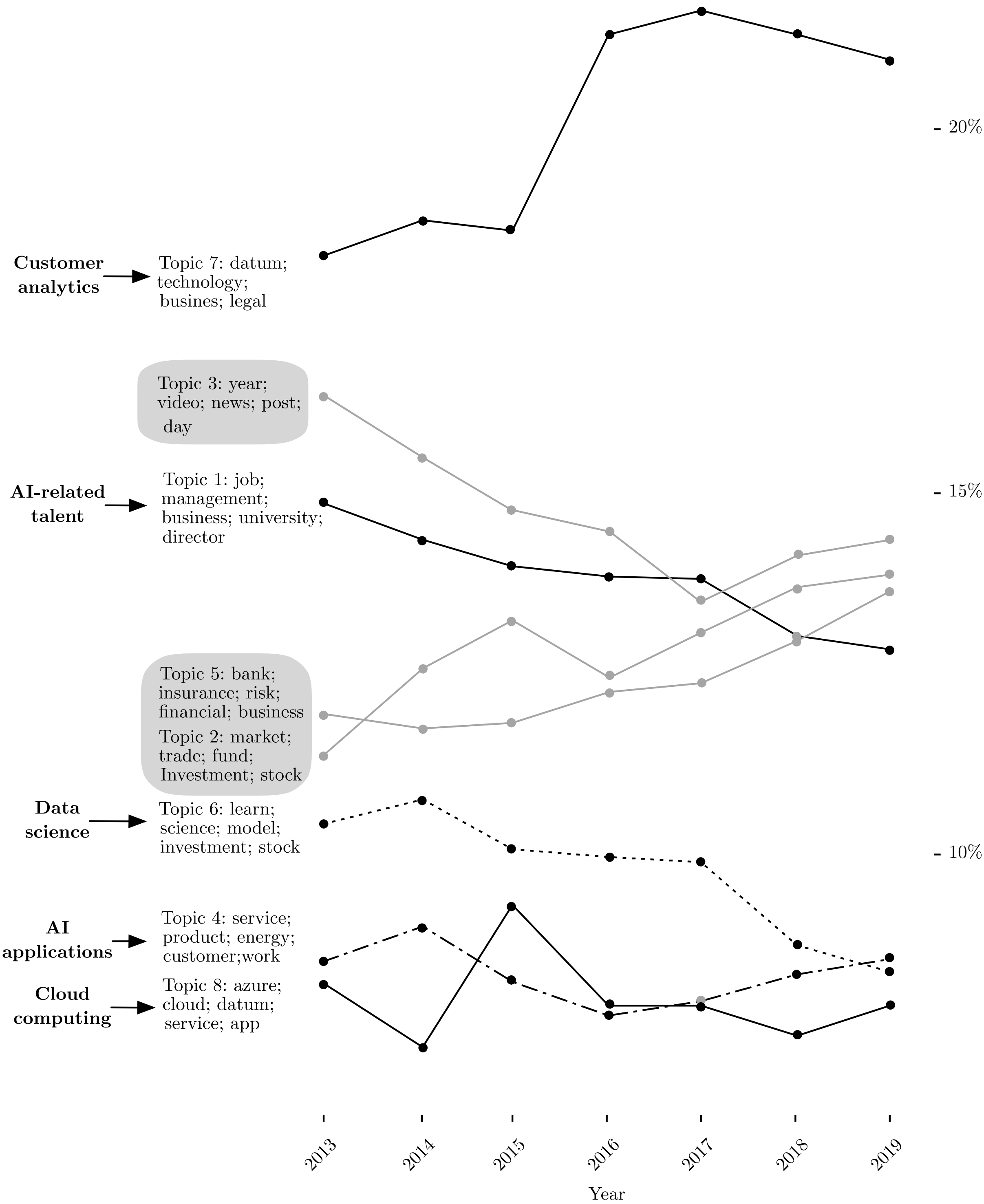
[§§] The retained documents’ length ranges between 5,000 and 50,000 characters. The retrieved web pages concern the largest 100 companies that operate in the financial service sector (as per SIC codes “64” and “65”) and are based in the City of London. Holding companies of corporate groups that do not focus on financial services are excluded from the sample.

Fig. 4. Saliency of the individual topics over time.***



*** Each data series indicates the likelihood that a document that is randomly drawn from the corpus will be associated with topic *i*. In the interest of interpretability, the far-right hand section of the chart illustrates the five topics that are most associated with each topic. Table 1 contains an expanded list of ten words per topic; shaded areas denote "general business context" topics.

Fig. 5. Saliency of the individual topics over time.[†††]



[†††] Each data series indicates the likelihood that a document that is randomly drawn from the corpus will be associated with topic *i*. In the interest of interpretability, the far-right hand section of the chart illustrates the five topics that are most associated with each topic; shaded areas denote “general business context” topics.

Fig. 6. Sequences of dominant topics across company and time.‡‡‡

69 seq. (n=69)

2013	2014	2015	2016	2017	2018	2019
5	1	3	5	5	5	5
5	3	5	1	3	8	7
3	1	1	3	1	3	3
2	4	2	1	3	4	4
1	1	3	7	7	7	7
3	2	2	2	2	2	2
1	2	2	2	2	2	2
3	3	3	3	6	3	3
1	1	1	1	7	7	7
7	7	1	7	7	7	7
3	3	5	3	5	5	5
1	7	1	1	2	2	2
3	3	3	3	7	3	3
3	3	3	7	3	7	7
8	5	2	1	1	2	2
3	5	1	7	5	5	5
1	3	3	1	1	1	1
3	3	3	3	3	3	3
2	2	2	2	2	2	2
3	6	6	6	7	2	7
5	8	5	5	1	8	2
1	1	1	3	3	3	1
5	5	5	5	5	5	5
8	2	8	2	1	2	2
1	1	1	1	1	7	1
8	4	1	3	1	1	1
6	6	6	6	6	6	6
7	5	3	7	7	7	7
7	7	7	7	7	7	7
7	7	7	7	7	7	7
7	7	7	7	7	7	7
2	2	2	1	1	1	1
1	4	8	8	1	2	2
2	7	7	1	3	7	7
3	1	3	3	7	7	3
2	2	2	2	1	2	2
2	2	2	2	2	2	2
2	2	2	2	2	2	2
1	3	2	2	2	2	2
2	2	8	2	2	2	2
8	4	7	7	4	4	4
3	4	1	4	5	5	5
3	1	7	7	7	2	7
2	2	2	7	2	2	2
3	7	7	6	7	7	7
8	1	8	4	4	4	4
7	7	7	7	7	7	7
1	7	4	1	1	1	1
2	2	2	2	2	2	2
2	2	1	2	2	2	1
3	3	3	3	3	3	3
3	3	3	3	3	3	3
3	4	3	3	1	3	3
1	2	2	2	2	2	2
2	2	2	2	2	2	2
1	7	3	7	7	4	7
1	1	5	8	5	1	3
1	5	3	7	7	7	7
5	2	3	7	7	7	7
7	7	7	7	7	7	7
7	3	7	7	7	7	7
3	7	7	5	7	7	7
5	5	5	5	5	5	5
2	2	2	2	2	2	2
5	4	3	5	3	5	5
5	7	7	7	7	5	7
1	3	5	3	6	6	3
3	1	3	5	5	5	5
3	1	1	2	2	7	1

Topics
1 – AI-related talent
2 – general business context
3 – general business context
4 – AI applications
5 – general business context
6 – data science
7 – customer analytics
8 – cloud computing

‡‡‡ For each company *i* and year *t*, we counted the number of times topic *i* has the highest probability to generate any document that belongs to the company-specific corpus of text *Di*. Then, we retained the topic with the highest count and considered it for the sequence analysis (conducted with R library TraMineR).

Table 1[*]

TABLE I
Term-Topic Matrix — 'Newspaper' Corpus

Term	Topics 1		2		3		4		5		6		7		8	
1	people	(0.011)	market	(0.021)	bank	(0.034)	datum	(0.033)	business	(0.02)	fund	(0.056)	company	(0.047)	firm	(0.044)
2	credit	(0.009)	rate	(0.014)	financial	(0.016)	company	(0.018)	bank	(0.018)	investor	(0.028)	venture	(0.025)	fund	(0.04)
3	good	(0.008)	rise	(0.013)	rule	(0.009)	technology	(0.013)	executive	(0.014)	price	(0.026)	startup	(0.016)	capital	(0.033)
4	loan	(0.008)	growth	(0.012)	regulator	(0.008)	customer	(0.012)	chief	(0.013)	investment	(0.026)	investor	(0.013)	company	(0.027)
5	thing	(0.007)	accord	(0.01)	firm	(0.007)	information	(0.01)	group	(0.012)	manager	(0.025)	lead	(0.013)	partner	(0.025)
6	risk	(0.007)	high	(0.009)	company	(0.007)	service	(0.01)	big	(0.012)	asset	(0.018)	technology	(0.012)	investment	(0.02)
7	question	(0.006)	fall	(0.009)	security	(0.006)	insurance	(0.009)	head	(0.01)	market	(0.018)	round	(0.012)	invest	(0.014)
8	pay	(0.006)	price	(0.009)	government	(0.005)	system	(0.009)	market	(0.009)	stock	(0.015)	include	(0.012)	accord	(0.013)
9	work	(0.006)	month	(0.008)	regulation	(0.004)	ai	(0.008)	industry	(0.008)	sell	(0.01)	capital	(0.011)	equity	(0.012)
10	lot	(0.006)	property	(0.008)	federal	(0.004)	insurer	(0.007)	client	(0.008)	management	(0.01)	funding	(0.011)	close	(0.012)

[*] Notes. — Labels for the topics: Topics # 1, 2, 3, 5, 6 and 8 = "general business context"; Topic #4 = "customer analytics (insurance)"; Topic #7 = "technology ventures"; estimations achieved with Mallet (McCallum, [2002]) software and the Gensim library for Python (Řehůřek & Sojka [2010]); number of documents = 5,036; number of topics = 8; terms are arranged in descending order of likelihood to appear in topic *i*; the optimal number of topics to retain is based on the comparison and contrast of the coherence value of 10 competing models in the 1-30 topics range — see Appendix A for further details about the estimation procedure.

Table 2†

Table II
Term-Topic Matrix — 'Internet' Corpus

Term	Topics: 1		2		3		4		5		6		7		8	
1	job	(0.012)	market	(0.03)	year	(0.007)	service	(0.017)	bank	(0.032)	learn	(0.011)	datum	(0.029)	azure	(0.045)
2	management	(0.012)	trade	(0.022)	video	(0.005)	product	(0.013)	insurance	(0.031)	science	(0.008)	technology	(0.014)	cloud	(0.026)
3	business	(0.011)	fund	(0.021)	news	(0.005)	energy	(0.012)	risk	(0.022)	model	(0.007)	analytics	(0.013)	datum	(0.021)
4	university	(0.01)	investment	(0.019)	post	(0.005)	news	(0.012)	financial	(0.011)	research	(0.007)	business	(0.012)	service	(0.018)
5	director	(0.01)	stock	(0.013)	day	(0.004)	technology	(0.012)	business	(0.011)	health	(0.007)	legal	(0.009)	app	(0.016)
6	work	(0.008)	investor	(0.011)	home	(0.004)	company	(0.011)	company	(0.01)	machine	(0.007)	customer	(0.009)	application	(0.012)
7	head	(0.007)	asset	(0.011)	work	(0.004)	business	(0.01)	claim	(0.009)	system	(0.006)	learn	(0.008)	build	(0.01)
8	group	(0.007)	global	(0.009)	world	(0.004)	health	(0.009)	service	(0.008)	datum	(0.006)	management	(0.007)	database	(0.009)
9	finance	(0.007)	price	(0.009)	share	(0.004)	medium	(0.009)	news	(0.007)	search	(0.005)	service	(0.007)	virtual	(0.009)
10	join	(0.007)	capital	(0.008)	medium	(0.004)	industry	(0.008)	group	(0.007)	law	(0.005)	work	(0.007)	scale	(0.009)

† Notes. — Labels for the topics: Topics # 2, 3, and 5 = "general business context"; Topic #1 = "AI-related talent"; Topic #4 = "AI applications"; Topic #6 = "data science"; Topic # 7 = "customer analytics (regulation)"; Topic # 8 = "cloud computing"; estimations achieved with Mallet (McCallum [2002]) software and the Gensim library for Python (Řehůřek & Sojka [2010]); number of documents = 21,612; number of topics = 8; terms are arranged in descending order of likelihood to appear in topic '*i*'; the optimal number of topics to retain is based on the comparison and contrast of the coherence value of 10 competing models in the 1-30 topics range — see Appendix B for further details about the estimation procedure.

Appendices

A — Business Press Corpus: data and methods

Data Gathering

- data source: Factiva
- targets: *Financial Times, The Economist, Wall Street Journal*
- details of the query:
  - timespan: 2000-Jan-01 - 2019-Mar-31
  - industry: financial services
  - region: all regions
  - language: English
  - duplicates: we discarded 'similar' articles
  - keywords: "artificial intelligence" or "deep learning" or "machine learning" or "big data" or "natural language processing" or "analytics"
- found: 5,301
- sampling:
  - *The Economist* has few articles (target removed)
  - before 2013, data are sparse (2000 - 2013 articles removed)
- retained for the analysis: 5,034 documents

Data Analysis

- articles passed through an NLP pipeline (spaCy) with the following characteristics:
  - tokenizer
  - entity recognizer
  - parser
  - English multi-task CNN model trained on OntoNotes, with GloVe vectors trained on Common Crawl (en_core_web_lg)
- pooled-cross sectional topic modeling/LDA (Gensim):
  - search over the solution space $s = \{5,6,7,8,9,10,15,25,30\}$
  - two models retained ($n_1 = 8$ and $n_2 = 30$) based on the distribution of coherence scores.

B — Internet Corpus based on City of London firms: data and methods

Sample of Companies

- legal & general assurance - aviva - lloyds - bank of england - standard chartered - prudential assurance company - royal mail - deloitte - valero energy - scottish widows - legal & general group - goldman sachs - aig - guardian royal exchange - old mutual public limited company - merril lynch - ashtead group - schroders - bupa insurance - axa - mace limited - investec - tp icap - isg plc - euroclear - blackrock - czarnikow - m & g group - ferroglobe - ubs - nomura – qbe insurance - carlin syndicate - mmc international - bank of america - bank of ireland – sumitomo mitsui banking - rbc europe - ivy holdco - willis limited - bechtel - brightsphere – intermediate capital - ig group - marsh & mclennan - group miki - hypersion - intercontinental exchange - xl insurance - hiscox dedicated - poundworld - talbot 2022 - acot underwriting - jefferies international - ardonagh midco - element materials technology - sainsbury's bank - beazley group - icbc standard bank - legal & general investment management - chubb capital - hcc international insurance company - munich re capital - novae corporate underwriting - mizhuo international - henderson investment funds - amtrust croporate member - ai mistral topco - aspen insurance - natwest covered bonds limited - howden uk group - hardy underwriting – scor uk - vtb capital - itau bba - investec wealth & investment - navigators corporate underwriters - tullett prebon - marex spectron - cna insurance - mfs international - pictet asset management - st andrew's insurance - citadel securities - q holdco limited - cmc markets - smbc nikko capital markets - cmc markets - aioi nissay dowa insurance - aberdeen asset management – ice future europe - franklin templeton investment management - ers corporate members - ig index - scotiabank europe - c. hoare & co - macquarie capital - anv corporate name - aetna insurance - westminster acquisition

Data Gathering

- data source: Internet
- targets: 100,000 web pages
- details of the query:
  - set of 100, City-based companies operating in financial services (holdings, listed under the same SIC codes of financial service companies, filtered out)
  - timespan: 2009-Jan-01 - 2019-Dec-31
  - region: all regions
  - language: English
  - keywords: “artificial intelligence" or “deep learning" or “machine learning" or “big data" or “natural language processing" or “analytics"
- found: 85,272 valid urls
- sampling:
  - file-filter: .docx, .pptx. .xlsx, .zip, .tar.gz, .rar, .7z, .pdf excluded. After the filter, there are 49,187 webpages
  - string-length-filter: webpages differ in terms of length. Topic modeling is sensitive to the distribution of the length of the individual documents in the corpus at hand (when there is substantial variation, the first topic extracted tends to correlate with the length of documents.) Based on the empirical distribution, we decided to retain documents with length s between 5,000 and 50,000 characters.
- retained for the analysis: 21,612 webpages.

Data Analysis

- articles passed through an NLP pipeline (spaCy) with the following characteristics:
  - tokenizer
  - entity recognizer
  - parser
  - English multi-task CNN model trained on OntoNotes, with GloVe vectors trained on Common Crawl (en_core_web_lg)
- pooled-cross sectional topic modeling/LDA (Gensim):
  - search over the solution space $s = \{5,6,7,8,9,10,15,25,30\}$
  - two models retained ($n_1 = 8$ and $n_2 = 25$) based on the distribution of coherence scores.