Excerpt from the original article “Artificial Intelligence’s New Clothes? From General Purpose Technology to Large Technical System”

Ekaterina Prytkova, PhD

Full article is published in the Journal of Information Technology.
Available at https://doi.org/10.1177/02683962231197824

AI System Components. Regardless of the methodological approach, any AI technology necessarily consists of the following components: (i) algorithms or virtual machines, (ii) computing power (and related physical devices delivering it), (iii) data, and (iv) domain structure as the problem environment and search space of actions an AI system is working with.

Challenges of AI ‘in the wild’: Data. We unpack the AI data domain to illustrate the incentives and bottlenecks it creates, the strategic challenges for AI development linked to it and, in general, the complexities related to AI deployment and value creation that involve **data**.

Over the last decade, we observe proliferation of business models that are reliant on data monetisation (and in particular Big Data). Proliferation and diffusion of digital technologies drastically lowered the related (information) search costs and the cost of tracking the consumption behaviour (content, goods, services, etc.) of online users. Atop of this abundance of data, new market opportunities for businesses that collect, store, structure, and elaborate the data grew rapidly: online databases, search engines, consulting firms, digital platforms, software management systems and many other examples of data-fuelled businesses. AI has the potential to spread into applications where (i) data is generated and can be collected in sufficient amounts and (ii) its structuring and elaboration creates value added for the business.

Step 1: Getting the Data. First, in order to deploy current AI to support any given application, an established and systematic process of data collection is required. In other words, the implementation of AI requires a good representation of business processes in data – namely, their digitisation. This is why pioneering industries in AI adoption have been the likes of Fintech and logistics, which are characterised by highly digitised and measurable processes and had forms of algorithmic automation and optimisation already in place. The current AI systems expanded the set of ‘digestible’ data by improving capabilities of processing of sensory data **at scale** such as images, video, audio, making activities that involve these capabilities cheaper (and thus economically viable), and less labour and time consuming. Thus, the current AI broadened horizons for automation.

The existence of data does not automatically make the case for an AI application. The collection of data that reflects business processes including demand’s feedback loops and the establishment of data markets is a necessary, though not sufficient, prerequisite for AI deployment.

Step 2: Choosing monetisation strategy. Second, to persist as a useful technology within an economic activity, data elaboration performed by AI has to bring returns. The value of data elaboration can lie in harnessing vast amounts and/or complexity of data detecting patterns in time infeasible for humans. Retrieving information about, for example, highly non-linear relations between a set of covariates and whether or not a person has clicked on an ad is undoubtedly a useful insight, but in order to systematically turn this information into profit a firm has to build a sustainable business model to monetize on it. Monetisation strategies can vary across applications, which in turn are characterised by different payoffs from the implementation of AI. For example, for online retail, the monetisation strategy would involve the structuring of pricing and versioning of the offer given the association revealed by data elaboration. This strategy allows obtaining profit directly and from each offer independently. Differently, an AI algorithm, that controls an industrial robot through the processing of sensory data and producing an adequate response in order to perform a task, creates value added that is more embedded and grows in a nonlinear way with the scale of deployment of the AI technology.

In sum, all kinds of data elaboration done by AI has to produce either valuable/unique result in the firm's production process or contribute to a valuable offer to the consumers, in both B2B and B2C markets, to ensure retention and generate profit.

Step 3: Investing in complementary assets. Third, the support of the monetisation strategy requires investments into complementary assets that constitute and/or support the AI system. In the case of data, the costs of primary collection or acquisition of data from third parties (e.g. the purchase of database licences, cookies or data appends), the storage within a firm or purchasing cloud space in order to further elaborate the data with AI, or even contracting micro-work to conduct data annotation, constitute yet another part of the data–related decision-making. Depending on the revenue stream from the activity that involves AI, a firm has to choose between investments into the development of AI systems at least in part in–house (including all domains – data, hardware and software) or in partnership with AI solutions providers at various levels of the AI value chain.

Small and medium-sized enterprises tend to gravitate towards outsourcing option to minimise costs. Even big companies for which AI performs a side function would be prone to purchase customised but ready-made AI solution, benefiting from sharing the risks and legal responsibilities with the developer. Indeed, among AI-users the emergent strategy of 'join-and-share' AI-as-a-service solutions due to the high costs of every component of AI systems steers AI development towards a form of infrastructure, with the most powerful actors (AI–producers) meticulously building and piecing together the AI stack.

Moreover, the burden of high costs is coupled with cross-domain network effects. For example, depending on the application, the nature of data might vary – pixel matrix for images, text corpus for legal disputes, or tabular data for consumer databases. This affects the choices and developments in the hardware domain (bandwidth capacity, memory size and placement, parallel or sequential processing and so on), programming framework (programming language, libraries), and algorithms themselves (loss function, optimization procedure). Together, the initial costs of implementation and cross-domain network effects increase the switching costs of an alternative to any component and lead quickly to hard lock-ins for both supply and demand in the software and hardware domains.

In sum, AI adopters make choices on how to deploy AI-based solutions and invest in the respective complementary assets.