**Data-driven innovation**


Data is the new oil. In 2006, mathematician Clive Humby used this analogy to represent the strategic need for companies to adopt a data-driven management process, in other words, to base all their decisions on careful analysis based on real data, making the decision-making process more structured and leaving empiricism behind. In the same year, Davenport published an article putting forward the idea that in the 21st century, competition on price or differentiation would become irrelevant, because only those that adopted analytical competition (use of data) would be able to generate sustainable competitive advantage. Almost two decades later, we see that data is perhaps even more valuable than oil. Widespread digitalization has reached people, companies and nations around the world, bringing connectivity and access to products, services and people simply and quickly. As a result of this process, millions of terabytes are generated every day, recording the steps, opinions, behaviors, habits and preferences of consumers and citizens around the world. Knowing how to extract and analyze this data opens up an unprecedented range of opportunities. At the same time, technological advances put pressure on organizations to innovate as a way of gaining or maintaining competitive advantage. Since innovation is a complex process that involves a great deal of uncertainty, could it be imagined as data-driven?

Data-driven innovation consists of extracting and mining large amounts of data (big data) from multiple sources to feed each phase of the innovation process. Data can be open and public or proprietary, can be generated by the user (e.g. social networks), by sensors (e.g. GPS), by companies (e.g. daily sales records) or by governments (e.g. open government databases) and can be structured (such as quantitative data) or unstructured (such as textual data, audio and images). Considering that the innovation process goes through the phases of opportunity identification, opportunity evaluation, solution designs and design evaluation, the data-driven approach can be incorporated as follows:

- Identifying opportunities: traditionally, in this phase, customer research is carried out, involving surveys, interviews, focus groups, in-store observation, A-B experiments, descriptive analysis of historical company data, in order to identify problems and unmet needs and thus an opportunity to innovate. In the data-driven approach, digital footprint data is analyzed using machine learning techniques, generally unsupervised (models in which response categories are not established in advance) with an essentially exploratory objective. An example might be the analysis of posts on social networks, resulting in a set of discussed topics that can provide insights into untapped markets, gaps in current products/services and, in general, stimulate the creative process.

- Opportunity assessment: traditionally, in this phase, experts try to translate the opportunities or needs identified in the previous phase into attributes that can guide the solution design process. The data-driven approach can enhance the work of these experts by analyzing customer preference data (based on browsing, likes or consumption) and identifying key attributes of the products and services purchased/targeted. With this, supervised models can automatically identify the key attributes for different customer preference patterns, making the work at this stage faster and better informed.

- Solution design: this creative ideation phase is traditionally based on the work of experts and brainstorming. The data-driven approach expands the design space by using generative algorithms (such as ChatGPT) to develop initial solutions based on databases of previous designs, which can then be improved by the experts.

- Design evaluation: in this phase an MVP (minimum viable product) is created which will undergo market testing with customers from the target audience. The prototyping and testing process can be slow and costly. The data-driven approach proposes an automatic test of a set of ideas, training a model with data from previous designs and their results (success/failure), so it is possible to discard ideas with a low probability of success before going into the field to test them with consumers.

In short, the processing of digital footprint data, customer preference data and data from the organization's internal knowledge is a powerful way of enriching the innovation process and possibly reducing or more effectively managing the uncertainty inherent in it. The reader may be thinking that the incorporation of data-driven innovation is specific to technology-based companies (big techs and start-ups) and/or large companies with big R&D&I budget. In fact, digital native companies have a culture of using data as a natural process, for innovation, CRM management and people management, among others. However, with today's technological landscape, implementing a data-driven culture is feasible even for smaller, non-digital native companies; a computer, cloud storage space, an analytics package or mastery of a programming language are enough resources to start a data-driven approach. Although the volume of data is high, the cost of obtaining and mining it is relatively low. In terms of human resources, for smaller companies it is not always possible to recruit a data scientist on the market to join the internal staff and start a data project, but it is feasible to train an internal member (there are dozens of good online programs on the subject) or even have the support of a specialized consultancy.

In any case, it's important that companies that need to innovate can and should surf the data age and use it to their advantage. Oil has never been available to anyone who wanted to extract it and benefit from it, data is.

Prof. Dr. Daielly Melina Nassif Mantovani

University of São Paulo - Department of Business Administration - FEA/USP

daielly@usp.br

## Further Reading

Jorzik, P., Klein, S. P., Kanbach, D. K., & Kraus, S. (2024). AI-driven business model innovation: A systematic review and research agenda. Journal of Business Research, 182, 114764. https://doi.org/10.1016/j.jbusres.2024.114764

Luo, J. (2023). Data-Driven Innovation: What Is It? IEEE Transactions on Engineering Management, 70(2), 784–790. https://doi.org/10.1109/TEM.2022.3145231