Scholedge International Journal of *Business Policy & Governance*
ISSN 2394-3351 Vol. 10, Issue 08 (2023)
DOI: https://dx.doi.org/10.19085/sijbpg100801

Could Synthetic Data Influence Strategic Thinking?

Prof Dr. MS S El Namaki
Dean, Victoria University, School of Management, Switzerland.
Dean (Retired) Maastricht School of Management, MSM, The Netherlands.

Abstract

Synthetic data is omni present. It is system produced and man-made. Their prevalence is induced by data gaps, data model analytics, system users and system abusers. They contribute, according to some estimates, a near 30% of applied data in AI analytical models today. This may grow by 2030, to a near 80% according to some sources. A development that is considered positive by some standards and less than positive by others. And the question is the extent to which this near avalanche of synthetic data could influence the strategic thinking of corporate bodies and, probably more importantly, their leaders.

This will be the core of the following article. The article is qualitative in approach. The ultimate goal is developing a set of hypotheses projecting a possible impact of synthetic data on the process of strategic thinking and the patterns of strategic behaviors in business. Analysis is based on current research in both synthetic data and strategic thinking. Synthetic data source material is fairly novel given the recency of the issue. Strategic thinking source material, on the other hand, relate to new developments in the concept. The ultimate outcome of the article should provide a base for research in two very relevant areas in current AI research: synthetic data and strategic thinking.

Keywords: synthetic data, AI, strategic management

1. What is synthetic data?

AI models need much more data and a greater variety than the real world can offer. Synthetic data provides a solution

Synthetic data is artificial data that is generated from original data. It could also take the shape of a model that is trained to reproduce the characteristics and structure of the original data. It is tantamount to the creation of artificial datasets simulating the original dataset's statistical characteristics. In particular, synthetic data allows for the generation of highly diverse or even novel data sets. It is also possible to enhance a training data set by generating synthetic data that does not reproduce the characteristics of the original data set, but instead exaggerates certain characteristics.

Synthetic data provides innovative solutions to problems of data scarcity and privacy, as well as algorithmic biases commonly, commonly used in machine learning applications. Synthetic data can be deployed to validate mathematical models and to train machine learning models. Data generated by a computer simulation can be seen as synthetic data. (Deng 2923).

Cite this article:
El Namaki, MSS. (2023). Could Synthetic Data Influence Strategic Thinking?. *Scholedge International Journal of Business Policy & Governance*, *10*(8), 80-86. https://dx.doi.org/10.19085/sijbpg100801

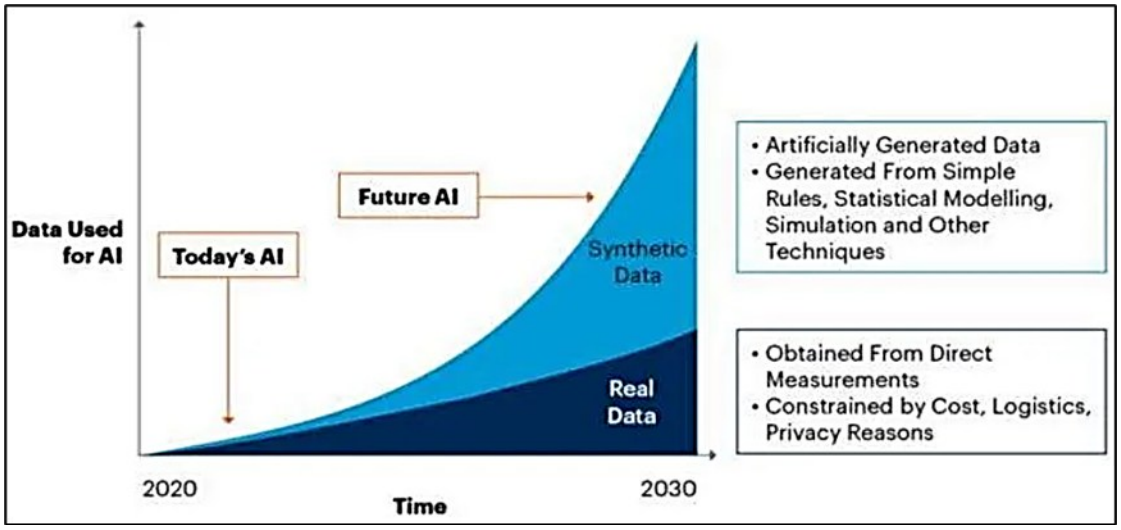

Figure 1: Contribution of synthetic data to total data population
Source: Gartner 750175

Generative AI, on the other hand, approaches synthetic data by examining the statistical distributions in a real dataset and creating a new, synthetic one to train other AI models. This synthetic 'pseudo' data is similar but not identical to the original, meaning it can also ensure privacy, skirt data regulations, and be freely shared or distributed.

How is synthetic data generated? Synthetic data can be generated using various methods. Those include:

- Generative Models or unsupervised learning task where a discovery and learning process leads to the discovery of the insights and patterns in real data in such a way that the model can be used to output new examples that match the same distribution as the real-world data it was trained on.
- Drawing numbers from a distribution or simply draw, or sample numbers from a distribution creating curve that is loosely based on real-world data
- Agent-Based Modeling (ABM) which is a simulation technique where individual agents people, cells, or even computer programs, are created that interact with each other.

Prime uses of synthetic data include privacy-by-design, safe data retention, testing software products and services, training ML and AI models, sharing data within organizations or without the organization, creating synthetic marketplaces and data exchanges and scaling out developments in the cloud. (Alex Watson, 2022), Data sensitive industries also turn to synthetic data in order to ensure privacy and regulatory compliance. Businesses are also leveraging synthetic data in domains where data availability is limited or privacy is a concern.

Synthetic data have their flaws. A key flaw is the "reality gap "or the subtle differences between the synthetic data and the real world. Another flaw is the benchmarking as synthetic data itself can also create (or even amplify) unintended biases. And yet another flaw could arise from Inferior quality of base data. Data quality could fall below the befitting standards of completeness, accuracy, reliability, relevance, and timeliness and could lead to synthetics that are equally inferior.

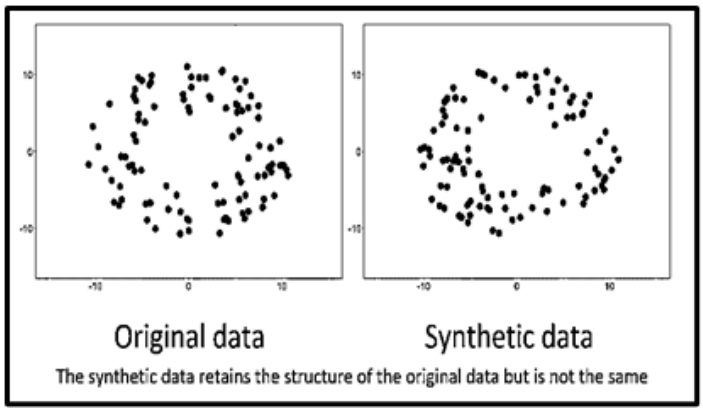

Figure 2: Synthetic data vs real data
Source: Synthetic data vs real data: which is a better option? (Sumit Singh, Labeler, Nov 15, 2022)

2. What is strategic thinking?

Strategic thinking is a cognitive process that delivers visions and goals and ways and means of their fulfillment.

Today's environment is both complex and uncertain. These conditions demand compatible business strategic thinking processes, allowing for perspective navigation of turbulent environments. Complex situations with shifting platforms require strategic insight and innovation. "In complex systems, one cause can create multiple effects. Reactive systems using previously learned behavior miss out on insight." (El Namaki, 2023). So, environments are evolving and the challenge of setting a strategy within an evolving environment is the need to understand that there is no single future but multiple futures, each with its own parameters and premises.

AI has provided a measure of response to the disrupted environment of strategic thinking. It has induced considerable shifts in the concept, the process, the vision, the insights, the competencies and the fulfillment efforts the strategic thinking process became an AI underlined system with inputs converted through strategic drivers into an output or fulfilled goals (Figure 2).

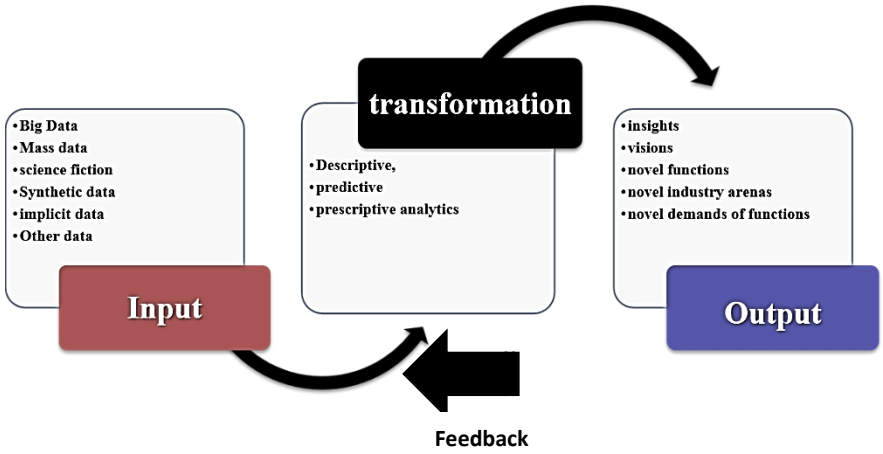

Figure 3: AI induced strategic thinking
Source: El Namaki, 2020

- **Inputs**

Data will constitute the prime input into future business strategic thinking models. This data will extend over the entire range from the real to the synthetic and from the big to the mass.

- **Transformation**

Transformation will, thanks to data analytics from the descriptive to the prescriptive and predictive analytics lead to the emergence of novel business arenas and derived instruments.

- **Outputs**

Output will include three varieties: sub system structures, novel functions and revealing insights. Insights might turn to be the most significant outcome here.

- **Feed back**

Feedback data will introduce essential system adjustments from inputs to transformation

This model departs from traditional business strategic thinking paradigms in several ways.

- First there is the trigger or the point of start of the strategic thinking process. Traditional analysis emphasizes an environmental scanning that could reveal "opportunities and threats". The suggested model's point of start is enhanced big data, or that mass of information with the wider scope and broader cover. Traditional scanning is replaced by databases that encompass almost every existing and possible variable of relevance to and of possible impact on this environment.
- Second, there is the conceiving of visions. Visions of the past were indeed based on a perception of futures to come. Visions of the enhanced big data era will be based on big data analysis and learning. A process that might allow these visions to go beyond the recognizable and familiar. They may address the unimaginable and the blurred. The distant and the far-reaching.
- Third, there are the tools of the analysis. Again, traditional analysis resorted to the common instruments within economic and social science domains. The above model suggests a reliance on advanced tools resting on a foundation of data analytics. Diagnostic analytics and predictive analytics will provide a strong impetus into the strategic thinking process. They will sketch a horizon that was unreachable before.
- Fourth is the earmarking of an arena or a field of business combat. The suggested paradigm leaves the door quite open to encounters that never happened before. Competitive encounters within uncharted arenas. Competition, in that sense, is replaced by either synergy or destruction by substitution.

- **Possible impact**

Synthetic data could have an impact on strategic thinking by influencing the drivers of strategic behavior or by altering firm strategic behavior itself.

- **The strategic driver's dimension**.

It is the authors contention that strategic thinking could be driven by either resources or technology. Capital constitutes the prime resource. Technology could represent science concepts and applications. Each is an independent force with a distinct force field and an equally distinct force impact.

Synthetic data could be viewed, as we mentioned above, as either gap filling or creative. Gap filling data sets are data based on existing data accessible within the existing real data conglomeration and is developed to fulfill model requirements. Synthetic creative data sets or newly created data sets or data "conglomerations" put together in order to fulfill so far unknown assumptions and, possibly, equally unknown premises.

The interaction between drivers of strategic thinking and synthetic data could lead to a variety of strategic spheres of business behavior or "business arenas":

- Gap filling synthetic data could relate to existing resource drivers strategic thinking and produce what we may term as "recognized business arena"
- Gap filling synthetic data could relate to technology driving forces and produce what we may term "expanded business arena"
- Innovative synthetic data could relate to existing resource drivers and produce "To be explored" business arena.
- Innovative synthetic data could relate to emerging technology driving forces and produce "unknown business arenas".

The most far reaching in terms of time event impact could be that where the technology driven strategic thinking pattern links with creative synthetic data sets. Both are movements into what one may call the "unknown". Technology driven strategic thinking derive its premises and direction from emerging technologies, an unknown sum that reveals itself as new concepts and relationships are revealed.

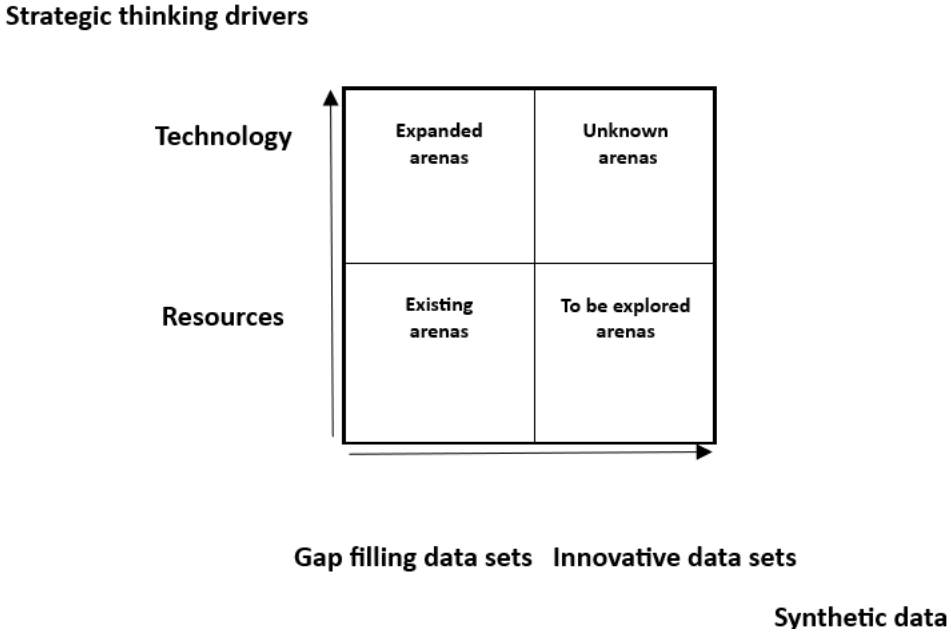

Figure 4: How synthetic data relate to strategic thinking drivers

- **The strategic behavior dimension**

Synthetic data could, on the other hand, relate to strategic behavior. Synthetic data could be viewed as we mentioned above as either gap filling or creative. Gap filling data sets are data based on existing data accessible within the existing data conglomeration. Creativity rooted data sets are data developed through access to broader data span and, possibly, inferred from distant data segments existing beyond the boundaries of the analytical model.

Strategic behavior here could reflect a pursual of the known or a pursual of the unknown. Knowns strategic behavior deals with existing products and markets. The unknowns in strategic behavior try to address emerging technologies and derived products and markets. The ultimate outcome is four situations.

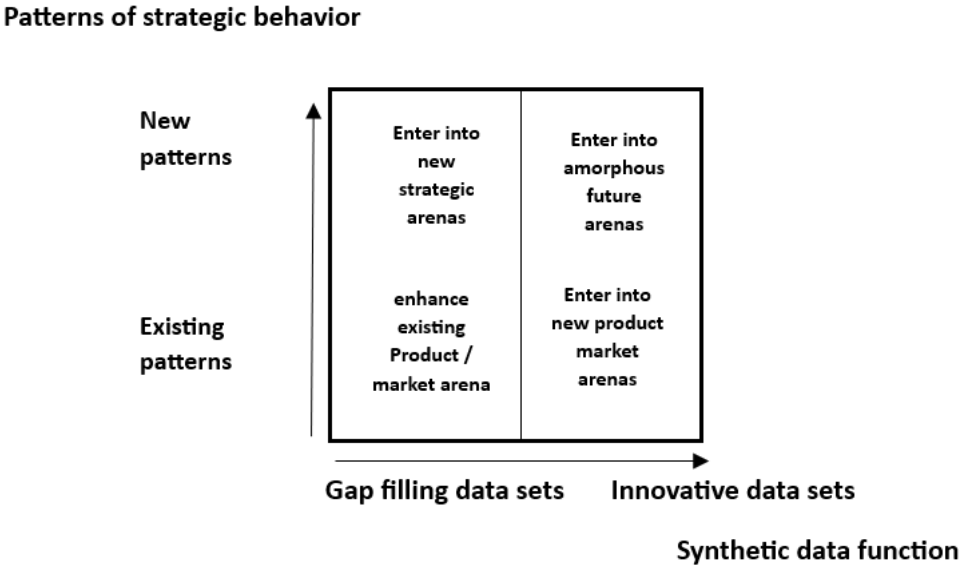

Figure 5: Relationship between synthetic data and patterns of strategic behavior

- A combination of existing patterns of strategic behavior and gap filling data sets could lead to enhanced existing product and market arenas
- A combination of new patterns of strategic behavior and gap filling synthetic data sets could lead to entry into new strategic arenas.
- A combination of innovative synthetic data and existing strategic thinking patterns could lead to entry into new product market arenas
- A combination of new patterns of strategic thinking and innovative data sets could lead to the entry into amorphous future arenas

Synthetic data would, in brief, lead to considerable change in the strategies and the strategic behavior of the respective business. The least intrusive is that where synthetic data plays a gap filling role and the most intrusive is when synthetic data assumes a creative profile and combines that with a pattern of strategic behavior seeking innovation and change.

3. Summary and Conclusions

Synthetic data is omni present. They are system produced and man-made. Their prevalence is induced by data gaps, data model analytics, system users and system abusers. They contribute, according to some estimates, a near 30% of applied data in AI analytical models today. This may grow by 2030, to a near 80% according to some sources. A development that is considered positive by some standards and less than positive by others. And the question is the extent to which this near avalanche of synthetic data could influence the strategic thinking of corporate bodies and, probably more importantly, their leaders.

The article is qualitative in approach. It sets a number of hypotheses projecting a possible impact of synthetic data on the process of strategic thinking and the patterns of strategic behaviors in business. The hypothesis relates to the strategic drivers as well as the strategic behavior resulting from resort to synthetic data.

Synthetic data would, in brief, lead to considerable change in the strategies and the strategic behavior of the respect business. The least intrusive is that where synthetic data plays a gap filling role, and the most intrusive is when synthetic data assumes a creative profile and combines that with a pattern of strategic behavior seeking innovation and change.

The outcome of this article will provide a base for derived research. This will possess considerable academic and applied value given current trends and the speed with which both issues are developing.

References

Deng, H. (2023). *Exploring Synthetic Data for Artificial Intelligence and Autonomous Systems: A Primer*, Geneva, Switzerland: UNIDIR, (2023).

El Namaki, M.S.S. (2023). Data-Driven Strategic Thinking. In: *Neo Strategic Management*. Palgrave Macmillan, Cham. https://doi.org/10.1007/978-3-031-37208-7_19.

El-Namaki, M. (2020). *From Product to Function: The Strategic Artificial Intelligence Shift*. Gartner, 750175.

Singh S. (2022). Synthetic data vs real data: which is a better option? , *Labeler*, Nov 15, 2022.

Watson, A. (2022). How to Generate Synthetic, Data: Tools and Techniques to Create Interchangeable Datasets, *Gretel*, March 22, 2022).

What is synthetic data? - Definition from WhatIs.com. *Search CIO*. Retrieved 2022-09-08).