# Note

Lucerne, 29 September 2025

**Roundtable: Large Language Models (LLMs) for market research – how useful are they?**

## The synthetic data journey so far

Synthetic data is explored in marketing campaign management, advertising, e-commerce, pharmaceutical processes, health, ESG/environmental data, and telecom.

- In marketing, synthetic "heavy buyers" have been generated to test campaign performance.
- In advertising, LLMs systematically better optimise ads on search engines than a panel of human experts. Instead of building a focus group, a feedback loop on words by an LLM seems to produce a robust optimisation.
- In e-commerce market research, where solutions like eye-tracking and product text optimisation are well established, the value of LLMs is still unclear.
- In pharma, laboratory processes are simulated to estimate duration and resource needs.
- In health, organisations simulate patient responses to treatments.
- Social and governance data are planned to be simulated to support sustainability assessments within the ESG (Environment Social and Governance) framework.
- In market research, an industry–academia collaboration tested synthetic personas, but results were only "ok" and not good enough for practical use. This raised the question of what kind of data is required.

## Successes and failures

Many applications of synthetic data appear as process optimisation for internal teams rather than market-facing products. Success stories are more often shared than failures. One example is a retailer–IT provider collaboration, where a simple configurator improved sales by 5–10% in one region. The AI handled the ideation, but the human factor was critical in deciding whether to adopt the suggestion or not. By contrast, failures — such as weak synthetic personas or "ok but not good enough" results — are acknowledged less openly.

## Challenges and Potential

A major challenge is evaluation. Blind spots often emerge only when models underperform in the field. The example of a large supermarket chain that struggled to generate useful personas shows that data quantity does not guarantee success. A similar failure was seen years ago in retail forecasting, where data was too heterogeneous for effective learning despite its abundance.

Another difficulty is that customer pain points do not always translate into willingness to pay for services. For instance, despite the frequent cases of counterfeit wine sales and stated concerns from both producers and consumers, there is little readiness to pay for wine authentication.

Prompt brittleness is another challenge everybody working with LLMs faces. This refers to the fact that minor changes to prompt formulation that are almost irrelevant for humans actually produce very different LLM outputs. This complicates the evaluation of LLM-produced synthetic data for marketing.

Future potential includes using synthetic data for market-facing products and processes, i.e. co-creation of services with customers. To illustrate, synthetic interviews were integrated into a design sprint in education, where the goal is to reduce project time from five to two days. The detailed analysis of a company's product range where each is analysed for competitors and niche opportunities also seems like a sweet spot for agentic LLMs.
Tools such as a digital workforce canvas or orchestration platforms like n8n provide convenient ways to set up and manage agent swarms.
In market research, synthetic data and LLMs should arguably not replace real data outright but instead maximise the effectiveness of human input. Feeding human feedback into the loop, akin to active learning, is considered promising. Ultimately, the challenge is to deliver insights and services that are fast, accurate, and credible.

**What should be reproduced?**
Synthetic data should imitate real data, which means it may also reproduce stereotypes or biases. Data should not be cleaned, or it risks losing realism. At the same time, care is needed to avoid discrimination, for example in loan decisions.

**Effective quality measurement and certification**
Both roundtable groups emphasised the need for evaluation at every step of the process, not just the end result. Drawing from systems engineering, stepwise checks such as unit, integration, and acceptance tests could help detect errors before they accumulate.
Certification and benchmarks were discussed as a possible way to guide model performance but was viewed critically as this may trigger reinforcement learning to meet the benchmark criterion. Yet key questions remain: who would issue the benchmark or certification, whether it should depend on the use case, and whether standards or processes might be more suitable than static benchmarks.

Questions and feedback should be addressed the roundtable organisers:
Simone Griesser, PhD: simone.griesser@hslu.ch
Simone Lionetti, PhD: simone.lionetti@hslu.ch