

# **Synthetic Data Generation: A Game-Changer in the Power and Utilities Sector**

**Dr. Bertalan Forstner**  
Emerging Technologies Strategist  
[bertalan.forstner@zenitech.co.uk](mailto:bertalan.forstner@zenitech.co.uk)

December 2024



# CONTENTS

---

Introduction .....	3
Why Synthetic Data Matters in Power and Utilities .....	6
The Business Benefits of Synthetic Data .....	7
Key Considerations for Implementing Synthetic Data .....	10
What Should Your Business Do Today? .....	11
Hype vs. Reality: What to Watch For .....	13
Final Thoughts .....	15





# Introduction

In an age of rapid digital transformation, the power and utilities sector faces unprecedented challenges—climate change, regulatory pressures, ageing infrastructure, and the growing need for cybersecurity are all at the forefront. As companies seek to innovate while managing risk and staying compliant, a key emerging trend is synthetic data generation. This technology has the potential to revolutionize how businesses operate, especially in high-risk, data-sensitive environments.

Synthetic data offers an opportunity to gain a competitive edge, without compromising on data privacy or quality. Let's explore why this trend matters and how it could address your pressing challenges.







# Why Synthetic Data Matters in Power and Utilities

Synthetic data refers to artificially generated datasets that mimic real-world data. This technology is gaining traction in sectors like autonomous driving, healthcare, and finance, but its relevance to power and utilities is just beginning to unfold. Here's why:

- **Regulatory Compliance.** With increasing regulations on environmental impact, synthetic data allows companies to simulate and monitor emissions and energy use in a controlled environment. This helps in preparing for audits and adhering to evolving standards without risking the exposure of sensitive operational data.
- **Ageing Infrastructure.** Companies can use synthetic data to simulate wear and tear on infrastructure, from power grids to pipelines. By running these simulations, businesses can anticipate and address failures before they become critical, extending the life of ageing assets and optimising capital investments.
- **Cybersecurity.** As operations become more digitised, the threat of cyberattacks looms larger. Synthetic data allows companies to test security protocols in a risk-free manner, identifying vulnerabilities in complex cyber-physical systems without exposing actual data.





# The Business Benefits of Synthetic Data

For forward-thinking leaders, **synthetic data generation** presents substantial and concrete business advantages. One significant benefit is **cost reduction**; acquiring real-world data can be costly, particularly in contexts involving physical infrastructure or sensitive environments. By utilising synthetic data, organisations can circumvent expensive field tests and effectively simulate various operational scenarios, thus minimising the need for physical interventions.

Another advantage is the **speed to market**. In today's fast-paced environment, rapid iteration is crucial for innovation. Synthetic data can accelerate this process by providing large and diverse datasets that facilitate quicker training of machine learning models. This enables companies to develop and deploy **AI-driven solutions** more swiftly, whether for applications in grid management, predictive maintenance, or energy storage optimisation.

Moreover, synthetic data enhances **privacy and compliance** in industries that manage sensitive information, such as customer usage patterns or infrastructure details. By leveraging artificial datasets, businesses can train models without exposing themselves to data breaches, thereby allowing for innovation while adhering to regulations like **GDPR**.





# Key Considerations for Implementing Synthetic Data

Before diving into **synthetic data generation**, it's essential to carefully evaluate several factors to maximize the benefits of this technology. First, the quality of the data, as synthetic data must be representative and of high quality, accurately reflecting the complexities and nuances of your real-world systems. Investing in robust synthetic data platforms that utilise advanced techniques like **Generative Adversarial Networks (GANs)** can help create reliable datasets. A robust dataset can withstand variations and still provide accurate insights; for example, a dataset representing multiple weather conditions and their impacts on energy consumption would be considered robust, while one that only reflects a single weather scenario is less reliable. The reliability of these sources can vary significantly; incorporating data from the most accurate sources ensures that the synthetic datasets generated reflect realistic and actionable scenarios. Also, when modelling synthetic data, leveraging historical data can provide essential context; for instance, analysing past weather patterns helps create more accurate energy consumption simulations under varying conditions. To maintain accuracy in your predictive models, regularly updating the synthetic data to align with real-time conditions, such as adjusting forecasts every 24 hours, can enhance the

effectiveness of AI-driven applications in grid management. By combining reliable historical data with updated forecasts, organizations can generate synthetic datasets that not only reflect past trends but also adapt to emerging patterns, ultimately supporting better decision-making in power and utility management.

Additionally, consider how **synthetic data** will integrate with your existing systems. It's vital to ensure that your data pipelines and **machine learning models** can seamlessly accommodate synthetic data alongside real-world datasets, as improper integration may lead to compatibility issues later on. To differentiate between proper and improper integration of synthetic data, assess how seamlessly the synthetic datasets enhance your existing machine-learning models without introducing biases or inaccuracies that could stem from conflicting data sources.

Lastly, weigh the **cost versus benefit** of implementing synthetic data technology. While it has the potential to reduce costs, developing or acquiring these platforms involves an initial investment. It's important to understand these costs and evaluate them against the potential long-term savings and advantages.



# What Should Your Business Do Today?

To stay ahead of the curve, power and utility companies should start experimenting with synthetic data generation now. First of all, we suggest that you **identify high-impact Use Cases**. Start by focusing on areas where synthetic data can deliver the most value. For example, testing AI models for predictive maintenance or running simulations for grid management. Should you get stuck during the process, **partner with our experts**. The technology landscape is evolving rapidly, and staying on top of the latest developments can be challenging. Partnering with our research-focused company allows you to tap into cutting-edge knowledge and stay competitive in the market.







# Hype vs. Reality: What to Watch For

Like any emerging technology, **synthetic data** comes with its share of hype, and it is important to recognise that not all solutions are equally effective. One significant concern is **overpromised accuracy**; synthetic data may not be as precise as advertised, so validating datasets and conducting rigorous testing before implementation in production environments is crucial. Additionally, while synthetic data is a powerful tool, it is **not a silver bullet**; some complex real-world phenomena still necessitate actual data for accurate modelling. Understanding the specific scenarios where synthetic data can be beneficial, as opposed to those where it falls short, is essential for maximizing its impact.

In the power and utilities sector, leading companies such as Siemens, IBM, and GE Digital currently dominate the software landscape. However, the applications of synthetic data remain largely untapped. Most existing solutions tend to emphasize **optimisation, grid management, and automation**, with few exploring the potential of synthetic data for **AI training** or **predictive controls**. This gap offers a significant opportunity for innovation within your business. By utilising synthetic data for advanced AI applications, you can surpass competitors who rely on limited real-world data, providing your company with a notable advantage in areas such as **predictive maintenance, cybersecurity, and infrastructure management**.



# Final Thoughts

Synthetic data is more than just a buzzword—it's a powerful tool that can transform how power and utility companies operate. For tech leaders, understanding and harnessing this technology now will provide an edge in solving tomorrow's challenges. Whether it's optimising grid management, enhancing cybersecurity, or preparing for the impact of climate change, synthetic data allows you to experiment, innovate, and stay ahead of the competition.

At Zenitech, we're not just looking at today's technologies; we're anticipating what's next. With our strong academic partnerships and cutting-edge research services, we're here to help you navigate the complexities of synthetic data and ensure that your business remains a leader in the industry. Contact us today to explore how we can help you implement this transformative technology and meet your business goals.