

Synthetic Data Services

G-Cloud 15





Table of Contents

1	Service Overview	3
1.1	'Synthetic' Data Generation	3
2	Business Need	3
3	Our Approach.....	3
4	Buyer Responsibilities	5
5	Service Management	5
6	Protection of Data	5
7	On-boarding and Off-boarding	5
8	Skills and Knowledge Transfer	6
9	Vendor Accreditations/Awards	6
10	Sub-contractors	6
11	Business Continuity and Disaster Recovery	6
12	Pricing	6
13	Ordering and Invoicing	7
14	Termination Terms	7
15	Further Information	7



1 Service Overview

1.1 'Synthetic' Data Generation

Using synthetically generated data supports organisations in reducing the risk of sensitive data leakage. Sogeti (part of Capgemini Group) have developed the Synthetic Data Generation solution - an Artificial Data Amplifier (ADA) which can generate synthetic data to be used to simulate live data in production and testing environments. The data can be generated "just-in-time" through DevOps pipelines or via a web-based user interface.

Sogeti's solution can generate a wide range of data types from structured (SQL, tabular), semi-structured (JSON, XML) to unstructured (images, text). With its containerised architecture, organisations can easily deploy and integrate the AI solution.

2 Business Need

The data requirements for modern applications are becoming increasingly complex, which means traditional methods of setting up test data through masking/obfuscation can become resource intensive, time consuming, and has risk of personal data leak.

Some of the business needs for synthetic data are as below:

1. Development and Testing
 - a. generate production quality data for higher grade of quality engineering.
 - b. generate testing data for ML models and accelerate the Dev and QA process.
 - c. generate data for performance testing.
2. GDPR Compliance
 - a. synthesize sensitive client information.
 - b. synthesize any other Personal Identifiable Information (PII).
 - c. make data sharing secure within organisation.
3. Need of more data
 - a. most AI implementations require large set of relevant sample data to train the AI model.
 - b. ADA can generate synthetic data to enable AI implementations.
4. Performance Testing
 - a. generate using AI will be able to maintain all characteristics of your data, improve testing capabilities and reduce the risk of sensitive data being leaked.

3 Our Approach

Sogeti's deep learning based Synthetic data generation solution, ADA, is the proposed solution for above mentioned business needs.

ADA is an AI solution that generates synthetic data that looks and feels like real data. It unlocks analytics and software development by supplying full access to data without compromising on customer trust, compliance, privacy and security.



Synthetic data generated using ADA is scalable, large amounts of data can be generated from a relatively small dataset. It uses advanced neural networks to map real data and generate synthetic data that preserves all the relationships & characteristics.

ADA can generate synthetic data in various forms like, tabular data (databases), Images and Free texts. Sogeti's approach to enabling organisations to utilise Artificial Data Amplifier (ADA) is described below -

5. Project initiation

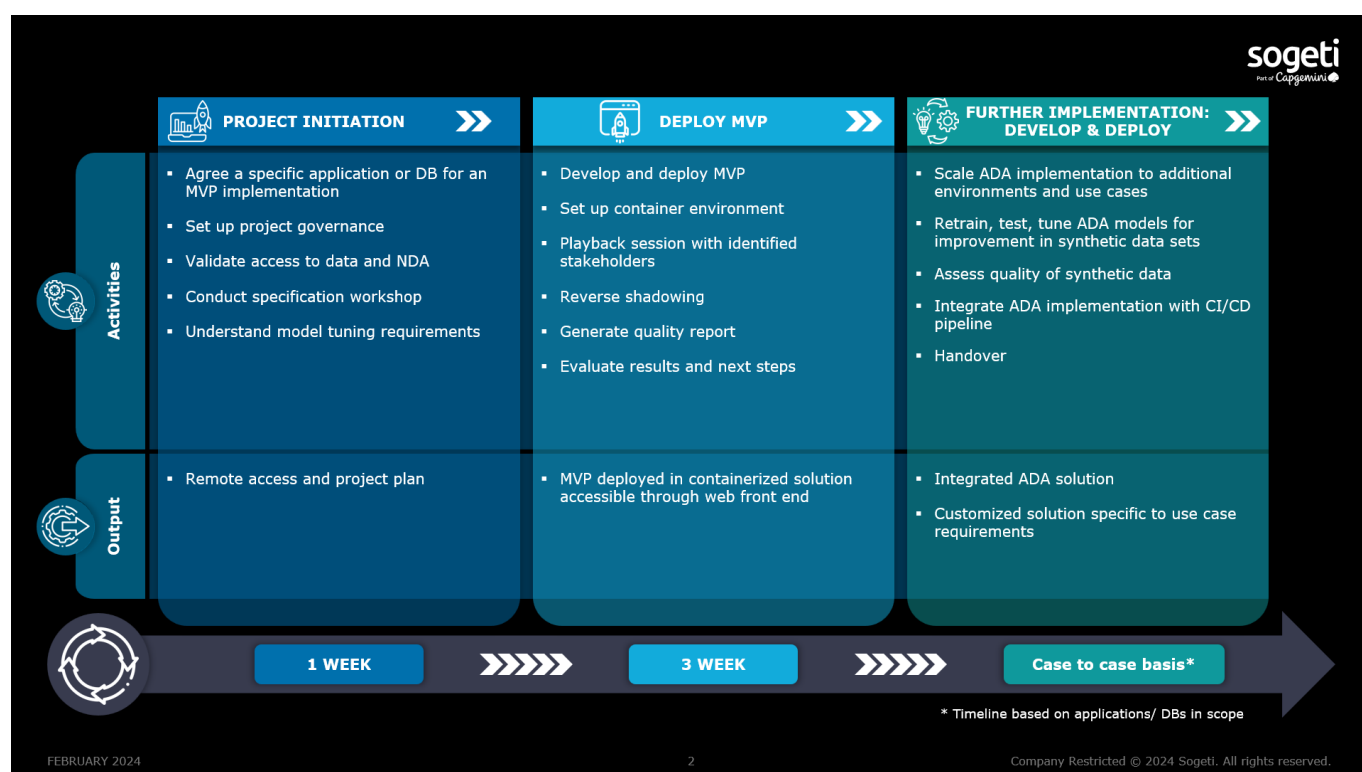
- a. Baseline application/DB for MVP use case development.
- b. Ensure access is available for consultants to develop the MVP.
- c. Agree project plan and governance.

6. Deploy MVP

- a. Develop MVP within a containerised environment that will be accessible through a web front end.
- b. Deploy the containerised MVP within client's environment.
- c. Collaborate with the client stakeholder's organisation to train and upskill key employees.
- d. Generate synthetic data and demonstrate data quality report.
- e. Conduct playback session with agreed client stakeholders.

7. Further implementation

- a. Scale ADA to other environments and use cases.
- b. Optimise ADA models by retraining, tuning, and testing to support bespoke requirements beyond MVP implementation.
- c. Handover to identified stakeholders / application teams.



This diagram is for illustration only and does not represent any obligation or responsibility of Sogeti.



Sogeti's synthetic data generation solution provides below validation mechanisms in form of reports -

1. Detectability Score

The detectability score serves as an indicator of whether your original data can be traced within your synthetic dataset.

2. Correlation Matrix

Correlation matrix score shows relationships between different variables or features of the original data with the synthetic data fields. It provides a comprehensive overview of how each variable correlates with every other variable.

3. Duplication Score

Duplication score indicates if there is presence of any original data in the synthetic data.

The data generated by Sogeti's 'Synthetic data support' provides value as below-

- Risk free and Scalable Data set – It enables GDPR compliant data, that enables the organisations of share the data within and outside organisation without any risks, as the data is synthetic and cannot be reverse engineered.
- Used for exponential opportunities – The dataset generated is of a quality that can be leveraged for various other purposes too. i.e., a Dataset is interchangeable with real data and can unlock & accelerate many complex AI solutions.
- High Quality and Speed – It can generate a high volume of data within relatively much less time.
- ETL and DB agnostic – the framework is ETL and DB agnostic, we can connect this to any database and infrastructure to train and generate synthetic data. like Oracle, SQL, SAP, MongoDB, Main frame etc.

4 Buyer Responsibilities

Please refer to the Supplier Terms listed with this service on the Platform. These may contain additional Buyer obligations/costs the Buyer is subject to that are not identified anywhere else in the Supplier's Application or on the Platform.

5 Service Management

This service can be delivered as a defined project or on a day rate basis.

6 Protection of Data

This service is based on a security classification of 'Official', however should you have a requirement for a different security classification that you would like us to consider, please contact us to discuss.

7 On-boarding and Off-boarding

Capgemini shall undertake on-boarding and off-boarding activities agreed within the Order Form (including as a minimum an exit plan in line with the Call-Off Contract terms) which will be charged for in accordance with the Pricing section for this service.



8 Skills and Knowledge Transfer

Capgemini recognises that skills and knowledge transfer is a critical element in the provision of GCloud services to public sector clients. Where possible and applicable, this forms part of the delivery plan for the service agreed at the start of the engagement. Our consultants and engineers are experienced in providing skills and knowledge transfer for major private and public sector clients.

Where appropriate, we may use a standard approach, tailored to topic, skills-gap and individual, to ensure consistency and effectiveness. The approach, Capgemini's Assess-Plan-Implement framework, has been used repeatedly by our teams to structure the work involved in transferring skills and creating new teams capable of driving and sustaining change long after the end of the formal programme. The framework can be applied throughout a project to understand knowledge transfer objectives, plan training delivery methods and materials, and deliver and evaluate success.

9 Vendor Accreditations/Awards

Sogeti as part of the Capgemini group, is an award-winning company both globally and nationally. The below list presents companies' names that Sogeti has won series of awards with over the years:



2025 winners of Best Advancing Software Testing Practice, awarded to the team or initiative that has made a positive contribution to the software testing profession. This is in recognition of those that go above and beyond to make the testing industry or practice better. It means breaking down barriers, thinking beyond our employers or clients, and using our skills and knowledge for the betterment of the profession.

Microsoft, IBM, Intel areas of Testing, HPE, Oracle, Digital, High Tech related awards, ESRI S-GROUP, Delivery, awards related to technologies and solutions.



For the 13th year in a row, Capgemini has been recognized as one of the World's Most Ethical Companies® by the Ethisphere® Institute. This is an acknowledgment of our ethical culture that makes us an employer of choice and responsible player in the eyes of our clients, shareholders, and the wider community.

10 Sub-contractors

Capgemini UK may use the following subcontractors to deliver this service:

- Capgemini Technology Services India Limited.

11 Business Continuity and Disaster Recovery

No disaster recovery plan is provided as part of these Services.

12 Pricing

This service is priced in accordance with the SFIA Rate Card attached. Capgemini can also provide offshore resources at reduced rates where appropriate. Projects can be priced either on a Time & Materials or Fixed Price basis.



13 Ordering and Invoicing

Please refer to the Supplier Terms for this service.

We would be pleased to arrange a call or meeting to discuss your requirements of our service in more detail.

14 Termination Terms

Please refer to the Supplier Terms for this service.

15 Further Information

For more information about this or any of our G-Cloud services, please contact our Public Sector Team.

Phone: 0370 904 4858

Email: publicsector.opps.uk@capgemini.com including the following information:

1. The name of this service.
2. The name of your organisation.
3. Your name and contact details.
4. A brief description of your business situation.
5. Your preferred timescales for starting the work.

About Capgemini

Capgemini is a global business and technology transformation partner, helping organizations to accelerate their dual transition to a digital and sustainable world, while creating tangible impact for enterprises and society. It is a responsible and diverse group of 340,000 team members in more than 50 countries. With its strong over 55-year heritage, Capgemini is trusted by its clients to unlock the value of technology to address the entire breadth of their business needs. It delivers end-to-end services and solutions leveraging strengths from strategy and design to engineering, all fueled by its market leading capabilities in AI, generative AI, cloud and data, combined with its deep industry expertise and partner ecosystem. The Group reported 2024 global revenues of €22.1 billion.

Make it real. | www.capgemini.com



This presentation contains information that may be privileged or confidential and is the property of the Capgemini Group.

Copyright © 2026 Capgemini. All rights reserved.