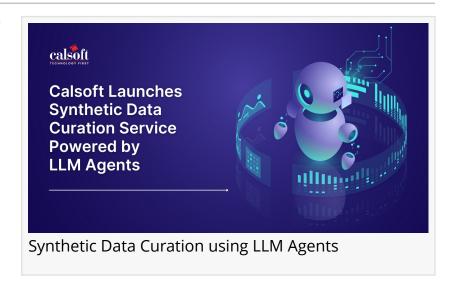


Calsoft Launches Synthetic Data Curation Service Powered by LLM Agents

Calsoft launches Synthetic Data Curation service powered by LLM agents, enabling enterprises to generate compliant, domain-specific datasets at scale for Al.



As enterprises accelerate Al adoption, many face a critical bottleneck: access to clean, contextual, and compliant data. Whether due to privacy concerns, limited labeled datasets, or the need to

"

We designed this solution to help teams move faster without compromising on compliance or quality" Ankur Somani, Associate VP – Technology at Calsoft simulate rare events, data gaps are slowing model development in key industries. Calsoft's new service addresses this challenge by enabling large-scale generation of high-quality synthetic datasets tailored to specific domains.

This service is particularly relevant for teams working on fine-tuning foundation models, building domain-specific classification systems, or testing AI behavior in edge-case

scenarios where real-world data is either insufficient or too sensitive to use. The ability to generate controlled, representative datasets without manual labeling or privacy concerns marks a shift in how enterprises can scale their Al pipelines.

 $000\ 00000000: 00-000000\ 0000\ 00000000\ 00\ 00000$

Enterprises struggling with scarce or restricted datasets can now leverage Calsoft's Al-driven curation service to:

- Scale: Leverage agentic pipelines of finetuned LLMs to generate millions of structured, semi-structured, or unstructured records within days
- Ensure Quality: Use multi-agent validation to maintain thematic accuracy, coherence, and domain alignment
- Maintain Compliance: Filter all outputs through PII-scrubbing and bias detection agents, ensuring governance at every step

Built on a closed-loop LLM agent architecture, the solution mirrors a human-in-the-loop process, where Generator agents create content, PII agents ensure privacy, Critic agents assess consistency, and Refiner agents improve flagged outputs.

In pilot programs with clients in finance and life sciences:

- Data preparation timelines were reduced by up to 70%
- Thematic accuracy in generated datasets exceeded 95%
- End-to-end delivery of productionready data took under 72 hours





These outcomes demonstrate how synthetic data can move from being a workaround to becoming a viable, production-ready alternative to real-world datasets. Teams that were previously slowed by access delays or redaction pipelines are now able to generate usable data in a fraction of the time.

"We designed this solution to help teams move faster without compromising on compliance or quality," said Ankur Somani, Associate VP – Technology at Calsoft. "With a closed-loop agent model, we're delivering scalable synthetic data pipelines that are technically sound and deployment-ready."

00000 00'0 00000 0000

This offering is being adopted across sectors such as:

- Regulated industries: Finance, Insurance, Healthcare, Life Sciences
- Data-intensive operations: Log analytics, eCommerce personalization
- Emerging use cases: Al education, experimental model testing

Organizations in these sectors face recurring challenges when it comes to acquiring and validating training data. With Calsoft's synthetic data curation service, these companies are able to accelerate time-to-value, reduce dependency on real user data, and extend model testing into scenarios that would be difficult to recreate otherwise.

"Our Synthetic Data Curation offering embodies Calsoft's vision to democratize AI," said Nilesh Chopda, Solution Architect. "By overcoming the data availability barrier, we're empowering enterprises to innovate responsibly and at scale."

With over two decades of experience in <u>digital product engineering</u> solutions, Calsoft focuses on technological innovation and engineering expertise to bring enterprise-grade safeguards to synthetic data curation, which includes:

- Peer-reviewed agent QA pipelines
- Built-in PII and bias filtering
- Custom domain adaptation
- Scalable delivery frameworks

These capabilities make the service suitable for enterprises looking to accelerate GenAl use cases while keeping control over data fidelity, risk, and compliance.

Calsoft is a global digital product engineering and Data & Al solutions company. For over two decades, it has partnered with enterprises and technology providers to accelerate product development, modernize platforms, and build Al-driven solutions across cloud, edge, and software-defined infrastructure.

Richa Thomas
Calsoft
+1 408-834-7086
email us here
Visit us on social media:
LinkedIn

This press release can be viewed online at: https://www.einpresswire.com/article/847462284

EIN Presswire's priority is source transparency. We do not allow opaque clients, and our editors try to be careful about weeding out false and misleading content. As a user, if you see something we have missed, please do bring it to our attention. Your help is welcome. EIN Presswire, Everyone's Internet News Presswire™, tries to define some of the boundaries that are reasonable in today's world. Please see our Editorial Guidelines for more information. © 1995-2025 Newsmatics Inc. All Right Reserved.