NextGenAI x air x University of Cambridge Judge Business School | Cambridge Centre for Alternative Finance
CAMBRIDGE SUPTECH LAB

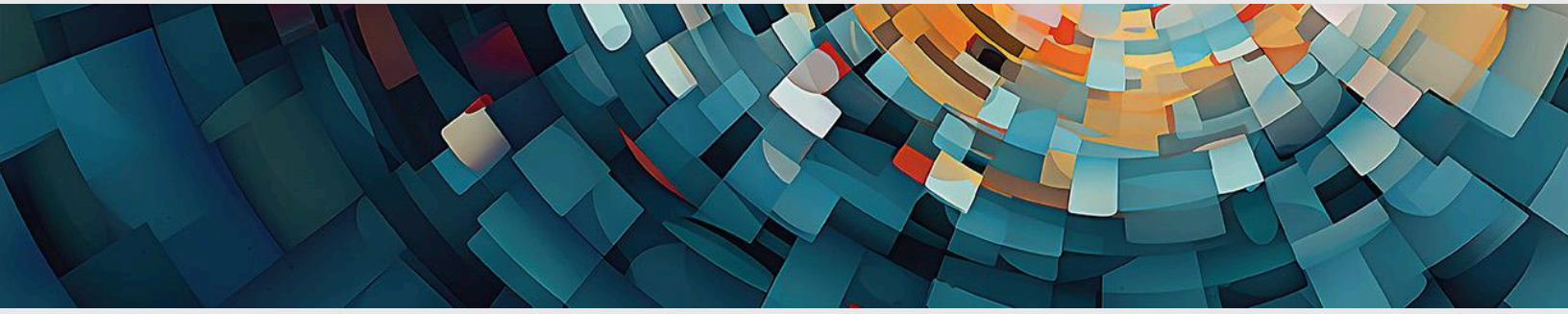
Key Learnings from NextGenAI Tech Showcase

The Alliance for Innovative Regulation's **NextGenAI Tech Showcase**, in collaboration with the Cambridge SupTech Lab, demonstrated various generative AI (Gen AI) enabled solutions over three sessions. Each session catered to different themes aligned with the **roundtable discussions** on the advantages and concerns surrounding the use of Gen AI in the realm of financial regulation. Following are the outcomes from the roundtables:

Advantages of using Gen AI

Financial regulators are keen on leveraging generative AI due to the following benefits:

- Increased productivity through automation of tasks like summarization, content creation, underwriting, investigations, and code generation
- Synthetic data creation to ensure data privacy and aid in bias reduction
- Virtual assistants and AI-based monitoring for improved customer service and compliance oversight
- Enhanced fraud detection and risk management through stress testing and scenario building
- Access to superior market intelligence, digitization of regulatory rulebooks, and accelerated data analysis to speed up the licensing process
- Personalization of financial services coupled with financial literacy tools contributing to consumer protection



Concerns around using Gen AI

While recognizing the advantages, regulators raised concerns regarding increased risk of hallucinations, the black box nature of algorithms leading to explainability issues, and potential biases in data, thereby underscoring the necessity for solutions catering to these downsides. Addressing these challenges involves implementing measures for human intervention, ensuring data accuracy, and enhancing transparency in algorithmic decision-making. Additionally, regulators are also focused on mitigating risks associated with data privacy breaches, phishing, and fraudulent schemes, emphasizing the need to improve oversight mechanisms to reduce fraud and protect consumers from harm.

How do the Tech Showcase solutions align with the roundtable outcomes?

The roundtable outcomes directly informed the focus areas and capabilities demonstrated by the various Gen AI enabled solutions at the Tech Showcase. The solutions aimed to leverage the advantages such as automation, synthetic data generation, development of virtual assistants and real-time monitoring, while addressing the key concerns raised by regulators around data privacy, oversight mechanisms, and potential misuse. The solutions demonstrated the following approaches to mitigate key concerns related to Gen AI adoption:

- Specialized language models coupled with retrieval-augmented architectures reduce hallucinations (incorrect outputs).
- Explainable AI techniques, human oversight, and audit trails address the "black box" nature and enable accountability.
- Expert-curated knowledge graphs, validation of AI outputs, and bias/quality checks tackle incorrect or biased data issues.
- Incorporating a ‘human-in-the-loop’ approach and leveraging domain expertise, including feedback loops and enabling iterative development involving regulators and subject matter experts.

The solutions spanned **themes** like data analysis, compliance automation, decision support, fraud detection, and consumer protection, aligning with the opportunities and risks discussed during the roundtables. This comprehensive approach showcased how Gen AI could be responsibly adopted by financial regulators while addressing their concerns through technical innovations and robust governance frameworks.

Theme-Wise Key Learnings

Theme 1: Data Analysis and Monitoring

Solution providers like Lingua Custodia, NayaOne, Complia, and Theta Lake showcased Gen AI capabilities for data analysis, information extraction and summarization, language translation, and synthetic data generation. Key learnings include:

- Specialized language models (combined with Retrieval Augmented Generation, RAG[1]) outperform general large language models[2] (LLMs) such as Generative Pre-trained Transformer[3] (GPT) for specific tasks like translation and domain-specific information extraction.
- Synthetic data generated via agent-based models or generative adversarial networks[4] (GANs) enables privacy-preserving data sharing, reduces bias, and allows risk-free testing across multiple use-cases and geographies.
- A conversational agent (mega-bot) placed on a network of AI agents[5] trained to cater to multiple use-cases involving regulations, privacy, financial crime, compliance, etc., can help with analysis, documentation and reporting across use-cases, thereby supporting the compliance monitoring and risk assessment processes.
- Some of these solutions can also be pre-trained and custom-built (be-spoke models) for end-users to enable very specific tasks.

[1] RAG - Technique that combines large language models (LLMs) with specialized knowledge databases
[2] LLM - Model trained on massive text data to understand and generate human-like text
[3] GPT - A type of LLM developed by Open AI
[4] GAN - Consists of two neural networks working against each other in an adversarial manner
[5] AI Agent - An autonomous entity that independently acts in line with its designed objectives

Theme 2: Compliance Automation

Solutions developed by Norm AI, Williams Lea, Braithwate, Cognitive View, Qatar Financial Centre Regulatory Authority (QFCRA), and Regxelerator demonstrated Gen AI capabilities for automating compliance activities, rules digitization, consumer compliance management, and supervisory assessments. Key learnings include:

- Some of these solutions use AI agents that act as autonomous decision makers and embody regulations, provide detailed compliance reviews, and enable scalable onboarding of new rules.
- Knowledge graphs[6] curated by experts and combined with transformer based ML/LLM models create digital twins[7], which provide outputs that enable traceability and accountability within a particular domain.
- Digital twins can also be used for production of machine-readable rules (i.e. rules digitization), followed by an assessment of impact of these rules for creating a more accurate policy making process.
- Gen AI models can assess risk by analyzing consumer complaints related to a regulated entity, identify compliance & conduct issues, and make recommendations with supporting evidence to ensure efficient consumer protection.
- Anti-money laundering (AML) compliance gap analysis, including compliance scoring, justification, and action points can be automated using vector stores[8] and LLMs.
- Gen AI and Robotic Process Automation (RPA)[9] can be used to create models that aid in the shift of the regulatory framework from a rigid rules-based approach to a principle-based approach which is more flexible and customizable.
- These solutions are being fine-tuned using a feedback loop process and can also be easily integrated with the other internal systems such as business intelligence or case management platforms using Application Programming Interface (APIs).[10]

---

[6] Knowledge graph - Structured representation of facts and relationships between entities
[7] Digital twin - Virtual replica of a physical object, process or system for simulation and analysis
[8] Vector store - Database to store and search high dimensional vector representations of data objects
[9] RPA - Refers to using software bots to automate repetitive digital tasks
[10] API - An interface that enables software interaction and data exchange

Theme 3: Decision Support

Solutions like Aptus AI, Winnow AI, DHI-AI, and Regnology are leveraging Gen AI capabilities for insights generation, information search, disclosure analysis, anomaly detection, and regulatory compliance support. These solutions are primarily designed in the form of dashboards, chatbots or virtual assistants. Key learnings include:

- Gen AI models fine-tuned on data can provide advisory, analysis, and decision support while reducing hallucinations by finding the right regulatory context. For example, custom LLM models trained on attorney-generated content enable accurate legal search and provide contextual information with supporting evidence.
- These solutions are using retrieval-augmented generation (RAG) architectures to answer questions and summarize findings, while providing references.
    - In the RAG architecture, domain specific data is available in a relational database which is integrated with the LLM to provide up-to-date, and contextually relevant responses.
- Multi-modal approach combining the strengths of LLMs (language understanding capability) and semantic search[11] (retrieval of relevant information) while leveraging human expertise in a feedback loop enhances accuracy of the output.
- Technology providers with solutions in the Minimal Viable Product (MVP) stage prefer to work closely with regulators using real-world data, and incorporate feedback from focus groups and subject matter experts to prioritize and refine features. (Eg: DHI-AI)

[11] Semantic search - Search technique that considers intent and contextual meaning, not just keywords

Theme 4: Fraud Detection

FeatureSpace, IDVerse, and FincrimeDynamics demonstrated Gen AI capabilities for transaction fraud detection, identity verification against deepfakes and synthetic fraud, and generation of high-fidelity synthetic data for financial crime simulations. Key learnings include:

- Large transactional models (LTMs) with behavioral embeddings[12] are used to predict future transactions and proactively identify fraudulent transactions, rather than solely relying on LLMs.
- Gen AI enhances transactional fraud detection models compared to traditional adaptive behavioral analytics, with an increase in fraud detection capability of 55% (vs. 33%) as stated by one of the solution providers.
- Regenerative AI[13] and deep neural networks[14] enable the detection of Gen AI-based identity fraud such as deepfakes, which is a growing concern in identity verification.
- Synthetic data generation combining Gen AI and agent-based models produces high-fidelity and privacy-enabled data, which helps in accurately simulating financial crime behavior.
    - This data is used to train the transactional monitoring models to detect fraud.
    - Benchmarking is done using ML models exhibiting similar behavior using both real and synthetic data to evaluate effectiveness of developed models.
- Scalability and global applicability of these models is a key differentiating factor.

[12] LTM with behavioral embedding - Gen AI model that analyzes patterns by embedding non-preferred behaviors within preferred ones

[13] Regenerative AI - Models that can self-improve and generate novel solutions iteratively

[14] Deep neural network - Artificial neural network with multiple hidden layers for complex data processing

Theme 5: Consumer Protection

Amplified Global and Proto showcased Gen AI capabilities for analyzing intelligibility and clarity of consumer terms and conditions, and developing multilingual customer support chatbots to ensure efficient complaint management, along with complaint analysis. Key learnings include:

- Gen AI can assess the intelligibility of consumer-facing communications, enable simplification, and provide audit trails for regulatory compliance.
    - Solution providers can ensure oversight by directly collaborating with regulators through their sandboxes and also by incorporating a human-in-loop approach to maintain accuracy and oversight at the client/end-user side.
- Multilingual chatbots using BERT[15] and GPT models enable customer support, complaint collection and information extraction, and sentiment analysis across various languages.
    - Additionally, regulators can provide access to the financial institutions (regulated entities) to check the complaints collected by chatbots and take appropriate actions.

Other solutions and themes: The Tech Showcase initiative also received solution submissions aligned with themes such as risk assessment (e.g., Elucidate) and policy development (e.g., Regology). Furthermore, solutions catering to the other prominent use-cases include marketing compliance (Haast), financial advisory compliance (Fifth Logic), and environmental, social and governance compliance (Deloitte). Further information on these solutions is available on the NextGenAI's **digital showcase** webpage.

For more information about AIR's NextGenAI Initiative visit regulationinnovation.org/nextgenai

[15] BERT - Bidirectional Encoder Representations from Transformers (Pre-trained LLM)