Tiger Analytics

# **Generative Al**

Tiger Analytics Capabilities

May 2024

### Our Generative AI Expertise Overview

### **Current Gen AI Capability Snapshot**



50+ Gen Al projects -Implemented & Ongoing



150+ Data scientists and MLFs with experience in Gen Al projects



Dedicated Gen Al R&D Lab



**60+** ongoing project discussions

### **Areas of Capability development**











Query to insights generation

Custom LLM model development

Intelligent search and summarization

Conversational agent

Content Generation

### Some of our current Gen Al client engagements

























#### **Notable Gen Al Industry Recognitions**

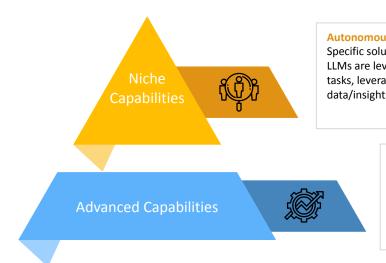


AIM's Top Gen AI Service **Providers 2023** 



Minsky Award for Excellence in AI 2023 Best AI Implementation by a Service **Provider** 

### LLMs are enabling a variety of Novel Experiences through these capabilities..



#### **Autonomous Agents**

Specific solution Agents built to execute goal/s; LLMs are leveraged to decompose goals, execute tasks, leverage tools or APIs and gather necessary data/insights.

#### **Next Gen. Chatbots**

Chatbots that can answer FAQs contextually, troubleshoot issues effectively, and even make small talks

#### **Query to Insights**

NLP-based conversational layer to analyze the data and respond to a question

#### Information Summarization

Intelligent AI and NLP based technology to understand queries and provide concise summary

#### Information Search

Insights from efficient information retrieval using NLP and Indexing

**Foundational Capabilities** 



#### **Content Generation**

Ability of foundational models to generate contextual & descriptive text from natural language prompts

#### Speech-to-text Generation

Ability of Generative Al model to understand and convert speech to text

#### **Code Generation**

Ability to foundational models to generate code from understanding natural language query

### Accelerating implementation of various use cases across industries

#### **Text Summarization**



#### **Text Categorization**



#### **Generate New Content**



#### Language Translation



**Generate Code** 



#### **Sentiment Analysis**



#### **Conversational Agent**

#### **Retail & CPG**

Summarize customer reviews, Sales reports, Social media outlook

Categorize products, Customer inquiries, Customer reviews

Product descriptions, Marketing emails, Social media posts

Multilingual customer service, Translate product descriptions, Customer reviews

Automate website updates, POS system updates, Inventory management system updates

Analyze customer feedback, Social media sentiment, Product reviews

Chatbots for Customer service, Sales, Feedback collection

#### **Banking & Insurance**

Summarize financial reports, Customer feedback, Regulatory compliance reports

Categorize transactions, Customer inquiries, Customer complaints

Customer emails, Marketing materials and reports

Multilingual banking services, Translate financial reports, Customer communications

Automate banking software updates, ATM software updates, Mobile banking app updates

Analyze customer feedback, Social media sentiment, Customer service interactions

Chatbots for Banking, Financial advice, Customer service

#### **Manufacturing**

Summarize production reports, Quality control data, Maintenance reports

Categorize parts,
Maintenance requests,
Safety incidents

Equipment manuals, Safety procedures, Product catalogs

Multilingual product manuals, Translate safety procedures, Production reports

Automate production line software, QC and Inventory Management software updates

Analyze employee feedback, Customer feedback, Supplier feedback

Chatbots for Production line troubleshooting, Maintenance scheduling, Supplier comms

#### **Health Care**

Summarize patient records, Medical research, Health outcome data

Categorize diseases, Patient inquiries, Medical research

Medical research articles,
Patient education materials,
Health newsletters

Multilingual patient communication, Translate medical records, Research articles

Automate health record system updates, Hospital mgmt. and Imaging software updates

Analyze patient feedback, Social media sentiment, Patient satisfaction Survey

Chatbots for Patient communication, Appointment scheduling, Medical advice

#### **Operations**

Summarize incident reports, Performance reports, Meeting minutes

Categorize tasks, Incident reports, Customer inquiries

Process documentation, Training materials, Internal communications

Multilingual operations management, Translate process documentation

Automate system updates, Workflow automation, IT infrastructure updates

Analyze employee feedback, Customer feedback, Supplier feedback

Chatbots for IT helpdesk, HR, Procurement

4

## Unlocking Value | Differentiators of Tiger's Generative AI Solutions



# Full Stack Gen Al Solutions

We build comprehensive solutions for our clients, covering data processing, model training, deployment, and UI/UX development, supported by a robust LLM governance framework



# Diverse and Comprehensive Skillset

With 150+ global experts in Data Science, Data and ML Engineering dedicated to Gen Al projects, we offer diverse expertise to craft holistic solutions, delivering significant value to clients.



# Cutting-edge Solution Accelerators

Our Gen Al solution accelerators empower clients to leverage LLMs for maximum impact, unlocking their data's full potential and revolutionizing business processes.



### Dedicated Research on Gen Al Tech

Our Gen Al Center of Excellence conducts ongoing research on topics like LLM Benchmarking, Guardrail evaluation, and Gen Al tools, improving solutions for clients.

#### **Wide Execution Spectrum**

Experience implementing diverse use cases, across domains, and complexities

#### **Partnership Ecosystem**

Extensive solution development with leading cloud providers: AWS, Azure & GCP

### Transparency / Open IP

White box approach with no licensing costs or proprietorship for our solutions

### Our experience with Gen AI and the impact we delivered to our clients

Retail

#### **Large Luxury Apparel and Accessories Retailer:**

Product Attribution Tool using Gen AI based CV model on Vertex AI that will help extract the product attribute, product title, product hierarchy and product description across 3 product categories

- Potential to reduce merchandising effort by more than 50% in year 1 through automation and streamlined collaboration
- Achieved extraction accuracy of 90% across product attributes for all 3 categories & 90% appropriateness of product titles

CPG

A multinational food, snack and beverage company: Developed a query engine with a realistic, talking human avatar-based user interface for the firm's R&D team that will enable users to query and search for data.

- 88% accuracy of transcription for the speech to text model
- Provides the ability to scale, connecting to various LLMs and third-party organizational directories, to multiple languages and to serve various departments within the organization

**Pharma** 

A multinational pharmaceutical company: Instant accurate information retrieval on Market Research, Biotech production and lab materials to enable better insights

- Built a **framework that can scale** to multiple business units, through specialized chatbots with minimal hallucinations.
- With 75% acceptance from SMEs, the solution answers user queries on market research documents with a response time of 9 seconds.

Call Center | BPO A global CX outsourcing leader with a network of 30K customers: Enable faster search for agents and retrieve relevant information to reduce turnaround time and improve call center efficiency

- 40% reduction in Average Handling Time of the agents
- Improved agent productivity generating 75% agent satisfaction at an accuracy of 80%
- Enhanced customer experience

### More Gen Al Success stories

1

Open AI based search and personalized product recommendation solution for a large American retailer 2

Leveraging Azure Open Al Chatbot for uncovering integrated insights around property search for a Real Estate Company 3

Gen Al Platform using
Insights Pro to understand
patient journey and
market performance of a
drug for a Leading
Pharmaceutical firm

4

Conversational Chatbot to enable information retrieval across 6 different engineering disciplines for a leading construction engineering firm

5

Integration of Custom LLM (Open AI) for generating interactive insights for a Large Health Insurance Company \_

Gen Al based Sales Agent Assistant for a Large Luxury Travel Retailer 5

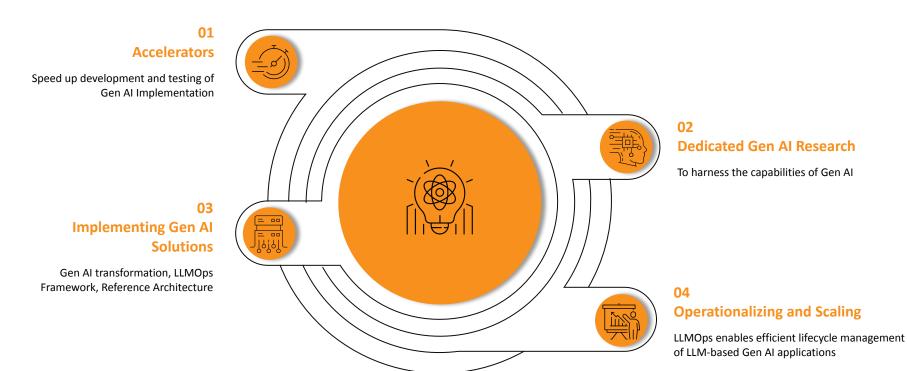
Evaluating LLMs across various cloud platforms to enable information retrieval for a Leading North American Manufacturer 8

Technology firm integrating Generative AI to track user statistics and create relevant dynamic dashboards

\_

Open AI based Knowledge Retrieval System for an Asian Government Agency

# We have invested in developing core capabilities enabled by LLMs / Gen Al..



## 1. Resulting in the creation of several Gen AI Accelerators

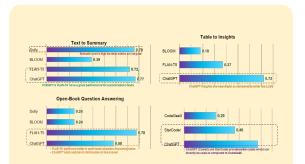
Retail & CPG	Banking & Financial Services	Insurance	Healthcare & Pharma	Manufacturing
Customer Service – Conversational agents for product recommendations & queries	<b>Lending –</b> Virtual assistant to address queries or generate insights on credit, risk	Pricing & Underwriting  -Analyze insurance history and enable decision making	Patient Journey — Summarize patient records to present accurate journey assessment	Supply Chain & Operations – Summarize production reports, QC data, maintenance reports for optimized planning & logistics
Marketing – Analyze customer feedback, Social media sentiment, Product reviews	Accounting — Summarize financial and Regulatory compliance reports	Claims – Reduce claims leakage through Intelligent search and summarization	Drug R&D — Analyze scientific papers, research articles and docs to generate insights and trends	Product R&D — Faster access to insights on engineering design and tools from a large document corpus

#### **LLMOps** – Development | Scaling & Deployment | Monitoring & Maintenance | Observability & Governance **BUSINESS ACCELERATORS Insights Pro Autonomous Agent Search & Summarization Consumer Insights Document Matching DATA ENGINEERING ENABLERS Spark Performance Tuning Assisted Data Modeling Code Generation** Log Analytics **Active Log Monitoring Manage Data Quality Intelligent Metadata Search Accelerated SDLC using Gen AI**

<sup>\*</sup>Use cases shown are representative/non-exhaustive list of problems Tiger's accelerators and enablers can address

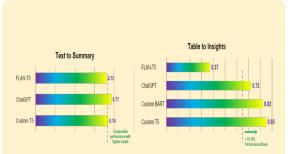
### 2. We continue to conduct dedicated Gen Al Research

#### **Benchmarking LLMs**



Comparing semantic similarity scores of LLMs Bloom, FLAN-T5, ChatGPT, Dolly across Text to Summary, Text to Insights and Open-Book Question Answering capabilities

#### **LLMs vs Custom LLMs**



Comparing LLMs and Custom LLMs performance on semantic similarity scores for FLAN-T5, ChatGPT, Custom T5, and Custom BART across Text to Summary, Table to Insights capabilities

#### Other areas of research



Evaluating various open source vector databases based on their compatibility with LLMs, and Enabling code-to-code transformations to optimize existing Python codes.

## 3. Getting prepared to benefit from the Gen AI transformation journey

A Gen AI Center of Excellence approach can help lay the foundation and accelerate the journey



#### **Awareness**

#### **Gen AI Data Discovery**

- Assess data availability and quality, technology readiness.
- Explore Data risks, PII and data exposures
- Bring industry expertise to define data foundation standards.

#### **Gen Al Playbook**

- Identify solutions and features that are best suited for Gen AI
- Framework for building an enterprise Gen AI Strategy
- · Emerging opportunities & feasibility study
- Assess impact and Generate an Action Plan



#### **Evaluation**

#### **Proof of Concept – Use case**

- Collaborate with business & data teams to Identify use-cases and technical resources for evaluating Gen AI capabilities
- Accelerate Al understanding within the organization & Develop an execution roadmap



#### Implementation

#### **Executing a Gen AI Roadmap**

- Develop and train AI models based on the client industry and specific context
- Co-own a roadmap & Provide Support and Advisory
- Agile roadmap to realize the vision

Advisory on the right model strategy to consider from PoC to Productionization

Off the shelf models

Open-source models tuned on demand

Build enterprise models grounds up

### 4. LLMOps – Operationalizing and Scaling Gen Al Solutions



#### Set of Tools and Best Practices - Manage the lifecycle of **LLM-based Gen AI applications**

Development, Deployment, Maintenance, and Governance of Applications Advantages:

- Efficiency faster go-to-market time
- Scalability lower operational costs
- · Secure and Agile Systems
- Strategic Decisions



#### What are the additional components compared to MLOps?

- Computation Resource
- Model Size and Complexity
- Efficient Inference & Latency
- Cost Effectiveness
- Prompt Management & Security
- Bias Mitigation
- Guardrails

**Data Science Team** 

(Experimentation and MLOps)

Develop the ML

models to solve

business problems

Human Feedback

#### **Challenges Addressed**

- Data Security/Leakage
- **Prompt Security**
- Secure Model Deployment
- Scaling Efficiency
- **Data Maturity**
- **Business Context**
- Silent Hallucinations
- Observability
- Reproducibility
- Responsible Al









Labeler Team

(Data Preparation at Scale)

Data Labelers/Editors

Label data for fine tuning

foundational models by

dedicated labeling tool

interacting with a

Advance Analytics Team

(Data Lake)

Managing Data Sharing and Provide Access

### ML Engineer

Collaborate with DS to productionalise the solutions

#### **Gen AI Team Extension** (Context Adaptation)

Select the corresponding foundational model, fine-tune and evaluate the model and design the deployment method/infrastructure

#### Platform Team (Cloud/Data/ML Platform)

Build CI/CD pipelines, user/service role, model consumption, testing and deployment methodology

#### Security/Architects

Assess data, user and services access creating policies and infrastructure

#### Risk and Compliance (Review and Approve Models)

Audit, Risk & Compliance Review models, Gen Al applications, data sources, code artifacts

Consumers of the Gen Al applications and rate the quality of output

Business (Viz Dashboard, ML

adoption, ROI)

Define business problems,

business KPIs and make

**Business Stakeholder** 

Consumers of Gen AI

applications, driving

**End Users (Consumer Gen AI** 

Applications)

**Generative AI End-users** 

business decision making

business decisions

**Business Consultants/Product** 

#### Gen Al Application Team (Gen Al Models in application)

Developer (Data Scientist)

Develop the back end of the

Gen Al application, Identify, evaluate and finalize the Gen Al tools and strategy to be used for the solution



Design the front end of the Gen Al application

Prompt Engineers

Design the input/outputs prompt to adopt the solution to the context and test the initial version

Prompt + UAT Testers

Test the Gen AI application at scale (back end/front end) and provide feedback for further improvement



### 4. LLM Governance | Tiger's Responsible AI Framework

#### Guardrails

Barriers that guide the ethical and responsible use of language models, ensuring adherence to established guidelines, preventing misuse, and promoting transparency and accountability in their deployment and decision-making processes

**Input Guardrails** 

**Prompt Guardrails** 

**Output Guardrails** 



#### **Evaluation**

Comprehensive assessment of LLM performance, capabilities and behavior across various metrics to provide a thorough understanding of the solution's strengths and limitations, guiding responsible deployment and usage

**Quantitative Metrics** 

**Qualitative Metrics** 

**Business Metrics** 

#### Monitoring

Framework to enable oversight of LLM implementations at model and application level across various metrics/ parameters

Model

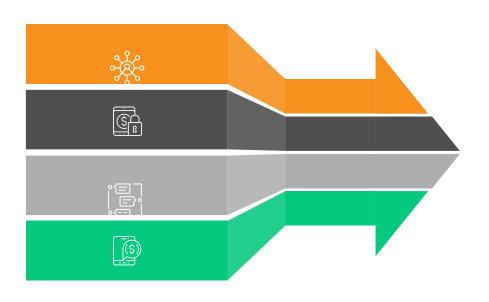
Application

User

### Gen AI – Reference Functional Architecture

#### **Digital Products** Application / Integration / Reporting Question Answering, Summarization, Conversational Agent, Output Search Engine, ChatBOT, RPA **Description Generation** Observability Monitoring & GuardRails, Bias & **LLMOps** Scaling & Deployment Logging & Testing (Prompt Tracking, **Ethical Responses** Maintenance Traces etc) LLM APIs - Open AI, Azure Prompt Engineering (CoT, Custom LLM (unsupervised, SFT Prompt tuning, Fine-tuning Modeling OpenAI, PaLM, BARD, Integration Zero-shot, one-shot etc) (LoRA, RLHF, PEFT) model, Reward Model, RLHF) **BLOOMBERG** ERP, HRMS, **Email** NLP Components -Intelligent **Embedding Store** Data Labelling (Prodigy, Knowledge Orchestration (Llama Pinecone, Weavite, PyAnno, Docanno, Extraction / Systems etc **Functional** Graphs (Neo4j, Standardization / Index, HayStack, fiass. Chroma Db. Humanloop, Scale AI), TigerGraph) Recognition LangChain) Qdrant Rule Engine Cloud Platform (Azure/AWS/GCP) Databricks/Others Infra Storage (ADLS/S3/Cloud Storage) Unstructured Data: Text data, Multimedia content (Images, Audio, Video files), Social Media feeds, Sensor and Telemetry Data, Webpage content, Emails, PDFs etc. Data **Structured Data**: Tables, Databases, Spreadsheets, CSV, JSON files etc.

### Our Engagement Models



#### Gen AI Center of Excellence

POD structure embedded into client ecosystem | Combined ownership of engagement and impact delivery

#### Fixed Price Delivery

Managed Program | Time bound engagements | Delivery focused

#### Accelerator led Delivery

Abstracted accelerators | white-boxed solution (no license fee) | Faster implementation and time to value | Premium Pricing

### Outcome based Delivery

New implementations | Unknown risks | Outcome owned by Tiger | Premium charged based on success

# In summary, a robust approach is required to lay a solid foundation for Gen Al led innovation

Goals **Business Impact** To ensure the outputs provided by Gen AI Robust outputs to drive confidence among **Comprehensive guardrails** overcome the known issues such as Silent users thereby increasing adoption Hallucinations Optimal use of To invoke LLMs only when necessary and Significant reduction in processing cost **Processing power** prioritize use of Frequently Cached and time Queries and Q&A Databases To support development team with the Ensuring the most up-to-date solution is **Dedicated** latest research, innovations and releases in deployed without any impact on release **Innovation Pod** Gen Al space schedule Iterative To frequently take user feedback to Allowing to deliver a user centric solution evaluate and improve insights that is easy to use and adapt Approach



#### **About Us**

Tiger Analytics is pioneering what AI and analytics can do to solve some of the toughest problems faced by organizations globally. We develop bespoke solutions powered by data and technology for several Fortune 500 companies. We have offices in multiple cities across the US, UK, India, and Singapore, and a substantial remote global workforce.

We have received multiple awards ranging from being recognized as a Leader by Forrester Research to being ranked among the fastest-growing tech companies by Inc. and Financial Times. We consistently feature in prestigious 'Best Analytics Firms' lists.

www.tigeranalytics.com







