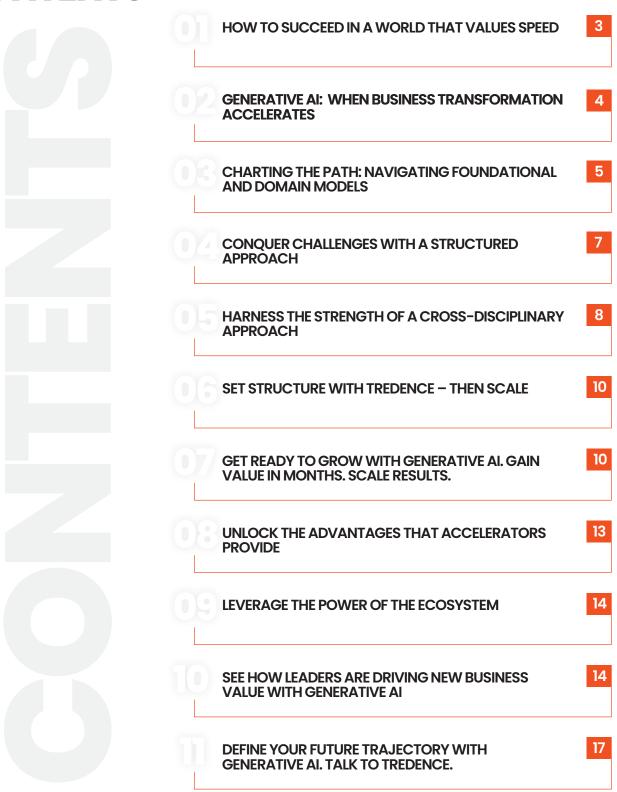
TREDENCE Beyond Possible

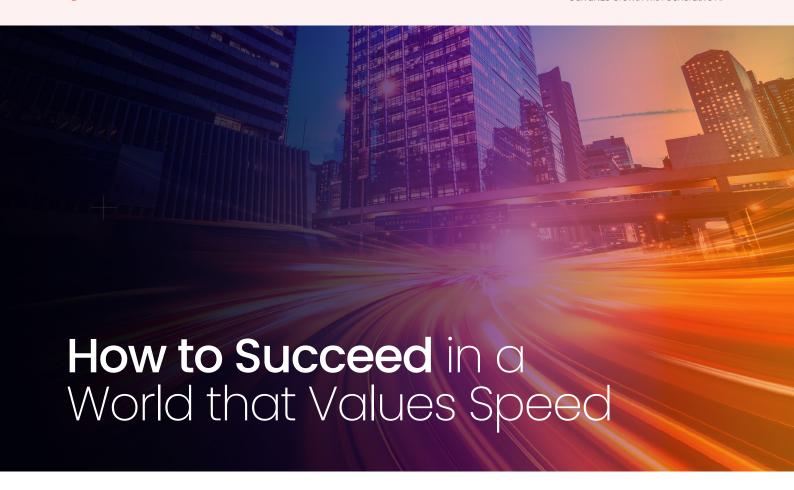
Generative Al Primer

# NAVIGATE GROWTH HORIZONS CENERATIVE A

Avoid Hype and Dashed Hopes. Unlock More Value with a Structure That Scales.

### TABLE OF CONTENTS





Generative AI is poised to drive workforce productivity and accelerate innovation worldwide. McKinsey projects that generative AI will automate 60 to 70 percent of employees' work today and unlock \$2.6 trillion to \$4.4 trillion in potential annual value across 63 high-value enterprise use cases.

Most of the world's businesses (70 percent) are exploring generative Al<sup>2</sup> to evaluate how this advanced technology can transform critical processes. Many will begin implementing new capabilities this year, while a few are already scaling them enterprise-wide. Those companies that deploy new generative Al capabilities at pace stand to create significant competitive advantage.

#### Get Ready to Galvanize Growth with Generative Al



Of generative Al's value will be reaped across four areas: optimizing customer operations, marketing and sales, software engineering, and R&D. <sup>3</sup>



Of generative AI high achievers have already deployed the technology across multiple business units. 4

- Michael Chui, Eric Hazan, Roger Roberts, Alex Singla, Kate Smaje, Alex Sukharevsky, Lareina Yee, and Rodney Zemmel, The economic potential of generative Al, report,
- 2 Agron Hurst, "70% of businesses currently exploring generative Al innovation," article, Information Age, May 3, 2023,
- 13 https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier#introduction
- 4 https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier#introduction

### **Generative AI:**

## When Business Transformation Accelerates

So, why is generative AI so critical to enterprise business transformation? Because this technology can understand, learn, and create, it provides a faster way to drive value with AI.

#### Generative AI enables enterprises and users to:

#### Do things

Enterprises can leverage generative AI tools, such as chatbots and virtual assistants, to boost worker productivity, automate repetitive tasks, and advance data analytics problem-solving. Teams gain by focusing on the work they love, reducing burnout and attrition.

#### Do things better

Companies can leverage partner lab and factory models, combined with comprehensive ownership, to enable rapid experimentation. Teams get to focus on innovating new capabilities rather than managing models.

#### Do better things

Teams can continually improve results by developing more targeted generative AI interventions. Customers gain by partnering with a forward-thinking enterprise that uses advanced technology in new ways to increase value. Shareholders benefit as companies drive ROI and reduce costs, increasing stock prices.

#### How Enterprises Are Using Generative Al Today

Companies have enabled use cases such as:



Deploying virtual assistants to enable better customer complaint management



Accelerating consumer product innovation and new formulations to drive ROI



Personalizing content to improve customer engagement and marketing campaign ROI



Summarizing complex industry, regulatory, or safety documents to improve worker effectiveness and accuracy



Orchestrating predictive maintenance processes to improve manufacturing yields



IT leaders and teams can choose among foundation and domain-specific models to achieve their objectives with generative Al. Partners can work with teams to determine which approach is best to solve the requirements of each use case.



#### **Building on foundation models**

Large-language models (LLMs) enable enterprises to build their own proprietary generative AI models in protected environments, including infrastructure, cloud, and data layers. Partners can provide intellectual property (IP) to train and finetune models, while clients use their IP to contextualize models to the use case further. LLMs provide a high degree of customization. However, this process is complex, time-consuming, and costly.



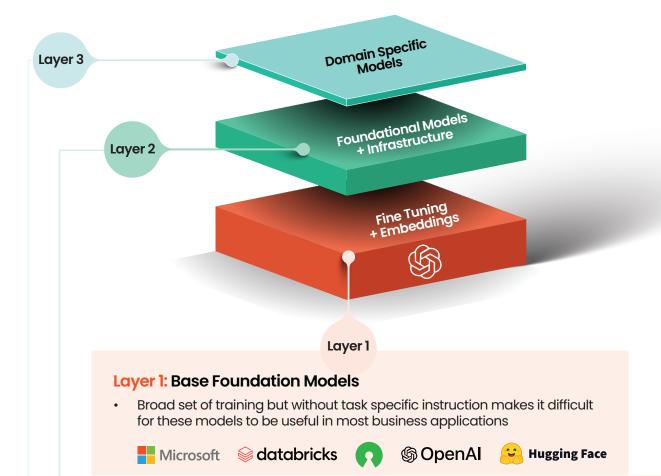
#### Leveraging partner domain-specific models

Partners can provide prebuilt, pre-trained models to speed time to value. Enterprise teams can develop embeddings and store the model in their protected environments. Partners can provide training and fine-tuning while teams use their IP to further contextualize the model to their use case. This process offers less flexibility but is simpler, faster, and cheaper.

As the graphic below indicates, foundation models provide general capabilities but require task-specific instruction and reinforcement learning to complete simple enterprise business tasks. Domain-specific models are trained on a specialized corpus, paired with extensive dialogue interactions or dialogue-annotated data, to execute higher-level tasks. This approach yields more accurate results and provides stable outputs across sessions and time.

# Focus of Most Business Applications

#### A Layered View of Applied GenAl Possibilities



#### Layer 2: Instruct/ Task Oriented Models

- Task specific instruction on human annotated data coupled with reinforcement learning provides more business relevant applications.
- The more similar or basic the task is versus the corpus of the foundation model, the lower the effort needed to customize.

#### Layer 3: Domain Specific Dialogue Oriented Models

- Domain specific instruction on specialized corpus often coupled with large amounts of prior dialogue interactions or dialogue annotated data to provide highly accurate and business relevant outputs
- Coupling with retrieval-based approaches and knowledge graphs increases relevancy, accuracy and stability of outputs

Caption: There is no one-size-fits-all approach with generative AI. Enterprises will use open-source and domain-specific models and tailor their approach based on use case business and technical requirements.

# Conquer Challenges with a Structured Approach

Beyond market hype and enterprise hope, challenges remain with adopting and scaling generative Al.

DIFFERENT CHALLENGES	HOW TO SOLVE THEM
Business Challenges	
Generative AI hallucinates answers.	Generative AI solutions can be designed to identify and remove erroneous content before it reaches users and approvers or is distributed to customers.
Models may demonstrate bias, creating business risks and harming users or customers.	Partners can provide responsible AI frameworks that build data privacy, fairness, and bias checks into initial data preparation, model development, and ongoing monitoring.
Generative AI can't ensure security and compliance.	Generative AI solutions can ensure security and compliance by leveraging a responsible AI framework, data segmentation, legal guardrails, and human intervention to identify and remove non-compliant content before it is distributed or used publicly.
Technical Challenges	
Generative AI content isn't ready to use out of the box.	Partners provide strategy, solution development, management services, and skill sets such as data science, data, prompt engineering, Al and model development, MLOps/LLMOps, and more.
Enterprises need to incorporate diverse data types and formats.	Partners can develop simplified integration frameworks that integrate various data inputs.
Skills and tools are evolving fast.	Partners offer leading-edge skills, develop and deploy proprietary algorithms, and stay current with open-source platforms.

© 2023 Tredence Inc. All Rights Reserved.

Operational Challenges	
LLM compute and API costs can soar if not well-managed.	Partners can provide effective Cloud FinOps capabilities to manage and reduce API and compute costs.
Teams must scope initiatives and demonstrate business value to maintain stakeholder buy-in.	Partners help develop business cases, prioritize use cases, share success stories, and provide metrics.
Teams will want to ensure that generative AI solutions are cost-effective and don't require vendor lock-in.	Enterprises can build, buy, or partner, depending on their requirements. Teams should ask partners if domain-specific models can be deployed as company IP to avoid ongoing licensing costs.
Initiatives that will result in significant process redesign require effective change management.	Partners leverage best practices to design initiatives, maintain stakeholder support, and train users on new tools.
Leaders and teams want to deliver value at scale.	Enterprises can leverage partners' lab and factory models and MLOps/LLMOps capabilities to enable generative AI experimentation and operationalize and productionize and monitor LLMs at scale.

# Harness the Strength of a Cross-Disciplinary Approach

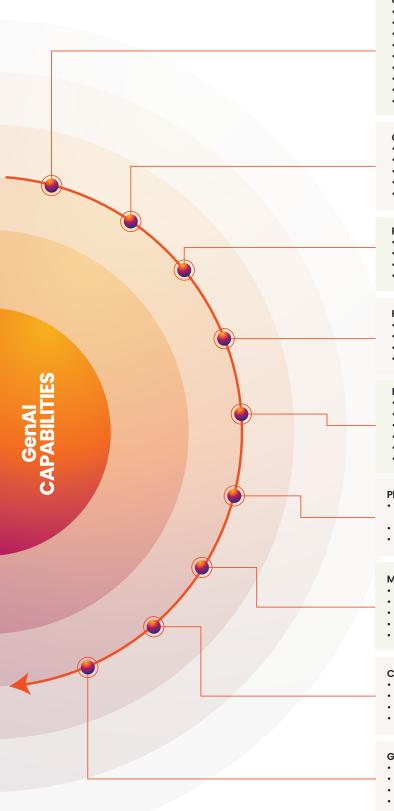
As generative challenges reveal, enterprises need multiple cross-disciplinary capabilities to succeed with this advanced technology.

Partners can provide end-to-end services, working with enterprise teams to develop an overall generative AI strategy and business case, create a technology strategy and roadmap to achieve these goals, and provide functional expertise to reimagine value chain processes and create a new business architecture and domain-specific corpus.

When it comes to, time to experiment, develop, and deploy, partners can provide UX capabilities, industrialized generative AI delivery capabilities, governance, and change management services. Enterprise teams can also tap managed services and platform engineering and implementation to scale model deployment and bring new capabilities to market at pace.

 $\hbox{@}$  2023 Tredence Inc. All Rights Reserved.

#### Range of cross-disciplinary GenAl capabilities necessary to develop, deploy, scale and drive adoption



#### GenAl Enterprise Strategy

- Portfolio approach/Roadmap Business Case
- **Use Cases**
- Performance Measurement & Monitoring Directions
- Financial Model/ROI
- Engagement Model
- Operating Model
  Alliances/Partnership Strategy
- Governance Approach and Ethics Strategy

- GenAl Tech Strategy
  Technology & Infra Choices (Build, Buy, Both)
- Technology Roadmap
- Platform Engineering & Automation Strategy
- Cost Models
- Alliances/Partnership

#### **Functional Expertise**

- Functional-Process value chain re-imagination
- **Process Maps**
- Business Architecture
- Domain specific corpus

#### Human-Centric Design

- Design Workshops
- User Journeys
- User Experience re-imagination

#### Industrialized GenAl Tech Delivery

- Lab + LLMCoE
- Infra + Assets + Frameworks
- Prompt Engineering Lib
- Prompt Engineering Templates Pre-Built Horizontal Vector DB store with Embeddings
- Pre-Build GenAl Solutions

- Platform Engineering & Implementation
  Integrated LangFlow Solutions, Multi-Model interaction frameworks, Templates, Automation Scripts
- Cognitive Search
- Cost Monitoring

#### **Managed Services**

- DataOps + MLOps + LLMOps
- Administration and LLM Management Workbenches
- SLA Management
- Continuous Monitoring and Continuous Improvement

#### Change Management & Adoption

- Program Management
- Stakeholder Management
- User Enablement & Training
- Evangelization

#### GenAl Governance & Compliance

- Policies Enforcement
- Standards, Security, Compliance Review
- Performance Measurement
- Reporting

# Set Structure with Tredence – then Scale

Tredence solves last-mile challenges with data, analytics, and AI, closing the gap between creating insights and unlocking their full value. We provide the expertise, services, tooling, and partnership enterprises need to succeed with generative AI. Tredence:



#### Provides full-spectrum generative AI skills

Tredence offers unrivaled expertise across the entire generative AI value chain, including LLM engineering, development, and operations; and platform engineering capabilities.



#### Offers domain and industry-specific expertise

Tredence combines data and AI strategy-setting and execution with customer analytics, marketing technology, supply chain, and other capabilities. We integrate these domain and industry insights to help enterprises crack the generative AI code, deploying new solutions that will unlock significant value.



#### **Delivers generative AI accelerators**

Our accelerators can be deployed within weeks, providing rapid results and helping enterprise teams create a culture of innovation and value discovery. We provide knowledge management, customer experience management, software development, and other accelerators.



#### Providing access to our ATOM.AI ecosystem

ATOM.AI, our intelligently engineered accelerator ecosystem, significantly reduces development time for generative AI pilot projects, enhancing efficiency and speed-to-market. We partner with hyperscalers to develop vertical solutions that speed time to value and speed to scale.

# Get Ready to Grow with Generative Al.

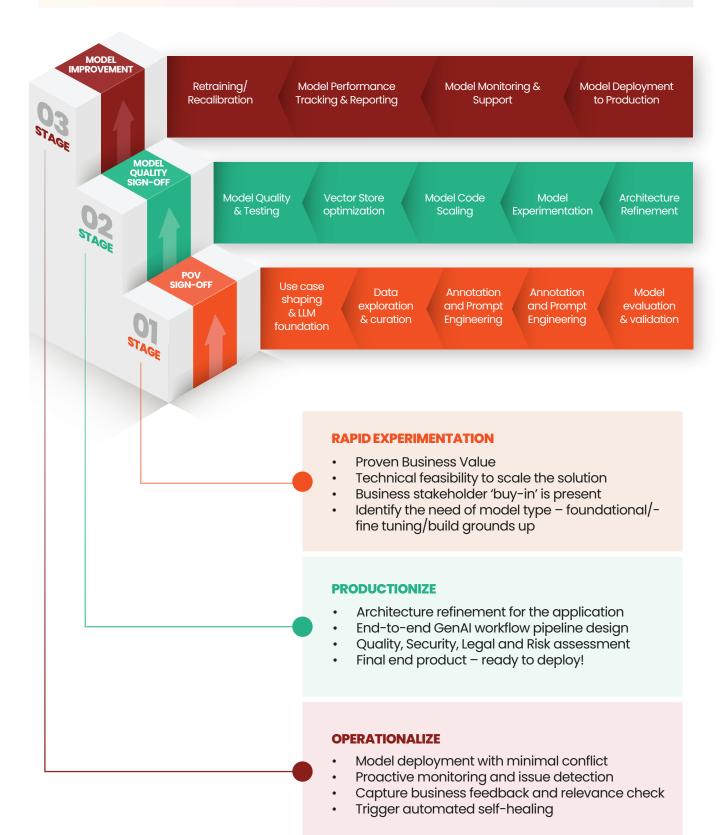
#### Gain Value in Months. Scale Results.

Drive workforce productivity and innovation with a generative AI strategy and plan that will unlock significant new value for your firm. By teaming with Tredence, you gain access to skills, technologies, and services to optimize the end-to-end generative AI model lifecycle, including experimentation, productionization, and operationalization.

During the experimentation phase, Tredence works with enterprises to select and develop a use case that provides rapid business value. During operationalization, we deploy and monitor the model, self-healing any drifts. And during productionization, we refine architectures, design pipelines, ensure security and compliance, and get the model ready to deploy.

#### **GenAl Ideation to Deployment**

**LLMOps** is **Dev Ops** tailored **for LLM** - a set of emerging best practices to enable significant improvements for industrialization of LLM systems



#### Your roadmap includes the following:



#### Discover your generative Al journey

Tredence will help you prioritize high-value use cases and perform a comprehensive gap-fit analysis to assess the current state and determine the necessary steps to bridge the gap between existing capabilities and desired outcomes. You gain a generative AI roadmap, customized architecture, and ROI-driven use cases within four to six weeks.



#### Develop a minimum viable product (MVP)

Next, Tredence will work with you to develop an MVP that will showcase generative Al's potential for transforming critical use cases, to provide rapid support and gain broad-based stakeholder support for the technology. This process includes using Figma to craft use-case designs; setting up cloud instances; and executing, deploying, and validating generative Al models. You gain a production-ready MVP within 10-12 weeks.



#### Scale your MVP

Tredence will work with your team to extend your MVP to broader business processes and integrate it enterprise-wide. We use best-in-class MLOps and LLMOps processes to enable continuous monitoring of models and optimize their performance. We also facilitate change management, allowing your team to transition easily to new processes.



#### Create a lab and factory model

To scale generative AI capabilities, you'll need to enable faster experimentation and learning. Tredence provides a lab and factory model with tools, analytics, and people; end-to-end ownership of generative AI capabilities from conception to production; and one-touch managed services. Our teams possess all relevant skills, such as data science and engineering, prompt engineering, DevOps, IT, MLOps, and LLMOps. We also offer accelerators to unlock new capabilities faster.



#### Planning for change management

Tredence provides end-to-end change and stakeholder management, from preparing business cases to implementing appropriate governance, training users, and working with teams to gain high adoption of new solutions.

# Unlock the Advantages that Accelerators Provide

Tredence provides several generative AI accelerators you can customize and deploy within weeks.

#### KNOWDY

#### Your employees' knowledgeable buddy

**(KNOW**ledgeable Bud**DY)** is an intelligent AI assistant that can be trained using your business-specific information, team details, workflows, and client data from your knowledge base securely.

#### **TREK**

#### Your organization's knowledge repository

**TReK** (Tredence Repository of Knowledge) is an Al-powered enterprise search solution that is a central hub for accessing essential documents related to various aspects of the organization. It provides a comprehensive reference and understanding of work, making project initiation easier.

#### COMPOGEN

#### Your no-code/low-code data science platform

A platform that combines descriptive, predictive, and prescriptive analytics into one interface, making data analysis and decision-making easier.

#### **ANALYTICS COPILOT**

#### Your data analytics, IntelliSense

Our GenAI technology transforms business analysis and data analytics, enhancing productivity and simplifying complex data interpretation.

Copilot empowers your enterprise, revitalizes your analytical framework, and gives you a competitive edge in today's evolving business landscape.

#### **GENFLOW**

#### Fastrack the development of GenAI/LLM accelerators

**GENFLOW** empowers developers, including non-coders, to quickly build valuable applications using large language models. It ensures visual appeal, operational efficiency, and continuous improvement.

# Leverage the Power of the Ecosystem

Tredence works with leading cloud, data, and AI partners to co-develop solutions for industry customers.











### See How Leaders Are Driving New Business Value with Generative Al

See how your peers leverage generative AI to solve common business challenges and drive more ROI.

#### Case study 1

#### Building GenAl Knowledge Management Tool and Establishing a Single Source of Truth for A Leading Investment Advisor

A leading American Investment Advisor required a Q&A Knowledge Management Tool for auditors. They aimed for a unified platform to assimilate data from varied sources and provide unambiguous answers to frequently asked questions. The tool will be accessible to the Internal Audit and Enterprise Risk team, featuring natural language query capabilities and text responses.

Tredence delivered an LLM-powered Knowledge Management Tool, implementing Retrieval Augmented Generation for coherent information synthesis and integrating data extraction from documents. The web-based interface, built on Vector DB, allowed auditors to submit queries and receive real-time pertinent answers. The entire GenAl knowledge management platform was built within a 14-week timeline.

#### **Results include:**



16-25% reduction in time-to-audit



**25-40**% reduction in report generation



25-35% improvement in auditor efficiency

#### Case study 2

#### **Enhancing Customer Experience for a Global Financial Services Firm**

A prominent financial services firm identified issues in their personalization and customer engagement journey, noticing that many customers would initiate the initial application process but frequently abandon it before completion.

Although the brand possessed substantial customer information, website personalization relied on rigid rules that did not fully harness the brand's available attributes. Messaging and communications were not customized to create a sense of the brand truly understanding its customers.

Utilizing the bank's composable CDP platform, we optimized a generative model foundation specifically tailored to the brand's standards. This model was created with the aim of accurately identifying the right messaging approach. Our goal was to harness the wealth of customer data and deliver personalized messaging, motivating them to successfully complete the application process.

#### The introduction of the personalization strategy had significant impacts:



**70%+** increase in new applications



**30%+** increase in completion

#### Case study 3

### Marine Classification Services Company Automate Regulatory Queries to Improve Field Services

This company classifies marine and offshore assets for customers. When surveyors in the field have questions, they contact subject matter experts (SMEs) for regulatory updates. However, routing and tracking challenges cause turnaround times of a few days to two weeks.

Tredence developed a generative AI query bot using Hugging Face and ChatGPT4 models. The bot automatically responds to routine questions but forwards them to SMEs when answers have a low confidence score. This approach could be used by any company seeking to improve knowledge management with regulatory issues and safety instructions.

80%

automation of all internal regulatory answers

**40**%

increase in survey or productivity

15-20<sub>%</sub>

improvement in classification accuracy

#### Case Study 4

### Enhancing Customer Issue Resolution and Categorization for a Retail Giant using Generative Al

The client faced challenges in effectively and accurately categorizing customer issues, resulting in delays and inconsistencies in problem resolution. They used driver transcript summaries with geographical, time, ride, and problem data. Calls were handled by agents who manually logged issues, potential solutions, and categorizations.



#### **GenAl Issue Summarization**

Tredence GenAl solution analyzed driver transcript summaries and automatically extracted and summarized customer issues. Automating this process led to a significant reduction in response time and an improvement in issue handling efficiency.



#### **AI-Driven Categorization**

The GenAl model was trained on historical data to understand patterns and context in customer-reported issues. GenAl was utilized to suggest the most appropriate categories and sub-categories for each reported issue, ensuring consistency and accuracy.

30%

reduction in issue resolution time



reduction in first response time

**15-20**%

increase in accuracy of issue resolution

#### Case study 5

### Ride Sharing Firm Streamlines Customer Complaint Response

A leading ride-sharing company has improved customer service by automating the categorization of customer complaints, reducing errors and time delays. Tredence implemented a generative AI summarization tool that analyzes driver transcript summaries, automatically extracting and synopsizing customer issues for call center agent review.



100% automated summarization of customer issues



20-25% increase in agent productivity



2-5% improvement in customer service

#### Case study 6

### Global Consumer Packaged Goods Firm Reverse-Engineers Competitor Product Formulations

A leading CPG firm is using generative AI to better understand competitors' product compositions and costs to increase the prices of its products.

Tredence has built a generative AI solution that uses publicly available data; leverages natural language processing techniques and a framework to extract and group ingredients; estimates the formulation and costs of products, sales prices, and gross margins; and adjusts the company's prices accordingly.



Unlocking competitor insights can reform product innovation, marketing and sales strategies, and more



\$5M: Each 1% price increase unlocks new revenues



Improves control over profitability

## Define Your Future Trajectory with Generative Al.

### Talk to Tredence.

You want to galvanize growth with generative AI: to develop leading capabilities that empower workers and deliver an exceptional customer experience. By partnering with Tredence, you gain access to the skill sets, frameworks, tools, and partnerships you need to unlock more value from generative AI. Deploy use cases, create a culture of experimentation, and equip business units and teams with new capabilities to innovate and execute at pace.

#### Seize This No-Risk Offer

#### Ready to get started?

Contact us to schedule a 60-minute discovery call, where we'll learn about your needs, discuss high-value use cases, and quantify the ROI you can achieve by using generative AI to transform critical workflows.

Outcompete by automating processes and improving decision-making with generative AI.

www.tredence.com/primer/an-extensive-gen-ai-primer

 $\hbox{\o}$  2023 Tredence Inc. All Rights Reserved.

#### **About Tredence Inc.**

Tredence is a global data science solutions provider focused on solving the last mile problem in Al. The 'last mile' is the gap between insight creation and value realization. Tredence is a Great Place to Work-Certified and as a 'Leader' in the Forrester Wave: Customer Analytics Services. Tredence is 2000 + employees strong with offices in San Jose, FosterCity, Chicago, London, Toronto, and Bangalore, with the largest companies in retail, CPG, hi-tech, telecom, healthcare, travel, and industrials as clients.







