

GenAI: Harnessing the Power and Responsibility

Texas State Agency Business Administrator's Association
(TSABAA) Mid-Winter Conference

January 18, 2024



What is Generative AI?

What is Generative AI?

A type of artificial intelligence that creates original content across various modalities (e.g., text, images, audio, code, voice, video) that would have previously taken human skill and expertise to create.

HOW does it work?

Uses **Large Language Models (LLMs)** trained on data, structured, semi-structured or unstructured, to **understand human communication and natural language**.

WHY now?

Converging innovations of hardware, cloud native stack, software engineering, machine learning, deep learning, and model sizes allows for the **realization of an autonomous creation economy**.

WHO is involved?

Technology leaders and start-ups are developing user facing applications on these underlying models.

- ✓ Act as a **force multiplier** for our people and our clients
- ✓ **Reduce the burden** of human interaction for many types of tasks
- ✓ Use technologies to **guide decisions** and **focus on critical tasks**

EXAMPLE MODALITIES

Text Generation

Prompt: *Explain my colleagues the business impact of generative AI in 50 words*

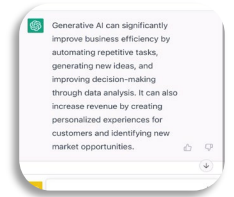


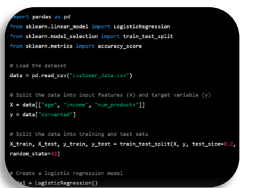
Image Generation

Prompt: *A bowl of soup that is a portal to another dimension as digital art*



Code Generation

Prompt: *In python, code a program that predicts the likelihood of customer conversion*



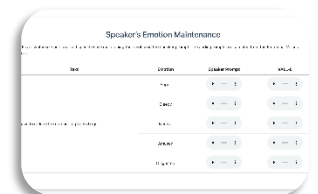
Video Generation

Prompt: *A teddy bear painting a portrait*



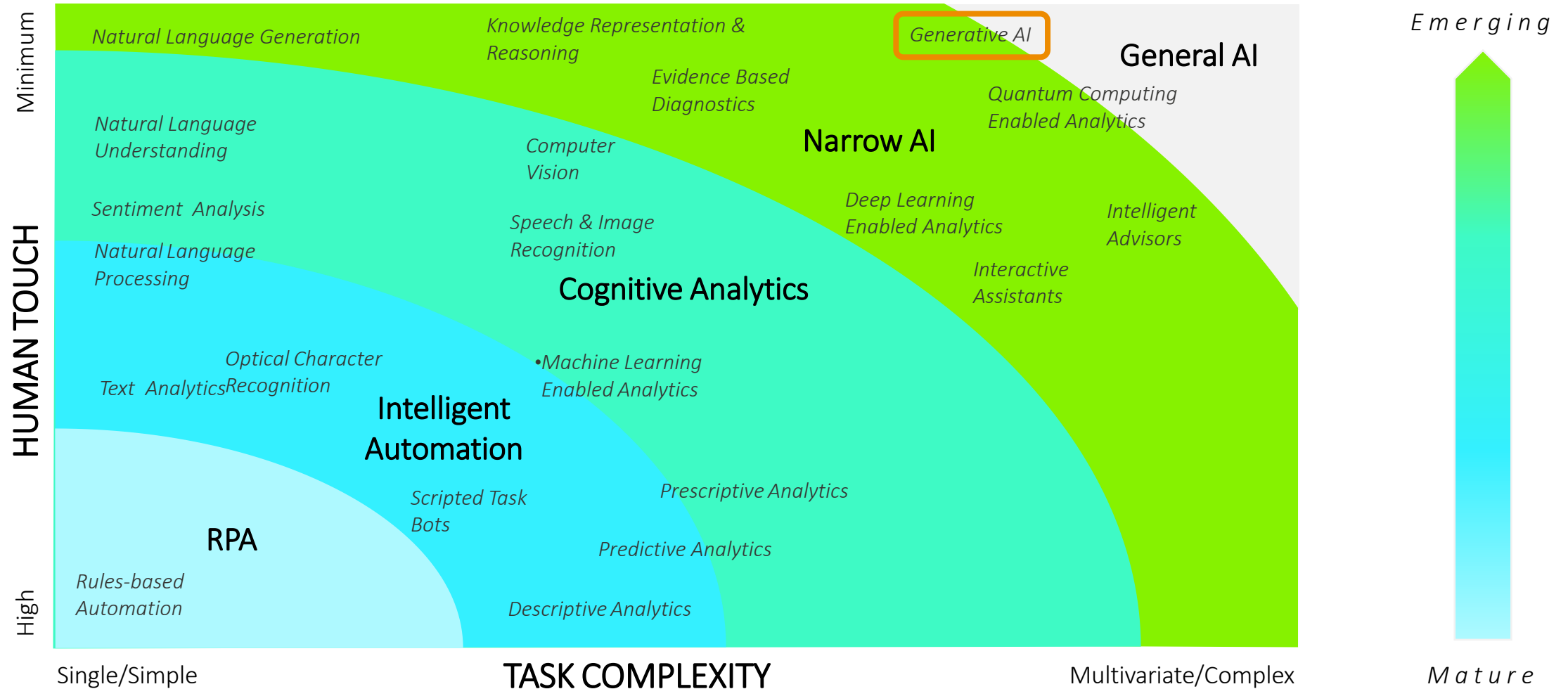
Audio Generation

Prompt: *Play 'we have to reduce the number of plastic bags' in a sleepy tone*



The Artificial Intelligence (AI) Spectrum

AI encompasses a wide range of technologies that use inputs of varying complexities to generate human-like outcomes.



Generative AI is One Facet of Artificial Intelligence

Organizations don't often request AI solutions by name, instead they want...

Cheaper Solutions

Inefficient use of capital and non-optimal decision-making wastes resources

Faster Workflows


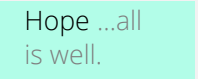







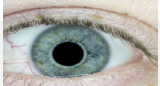




Mundane, repetitive tasks take away from mission-critical work. Speed to action is stalled in manual review and analysis

Better Insights

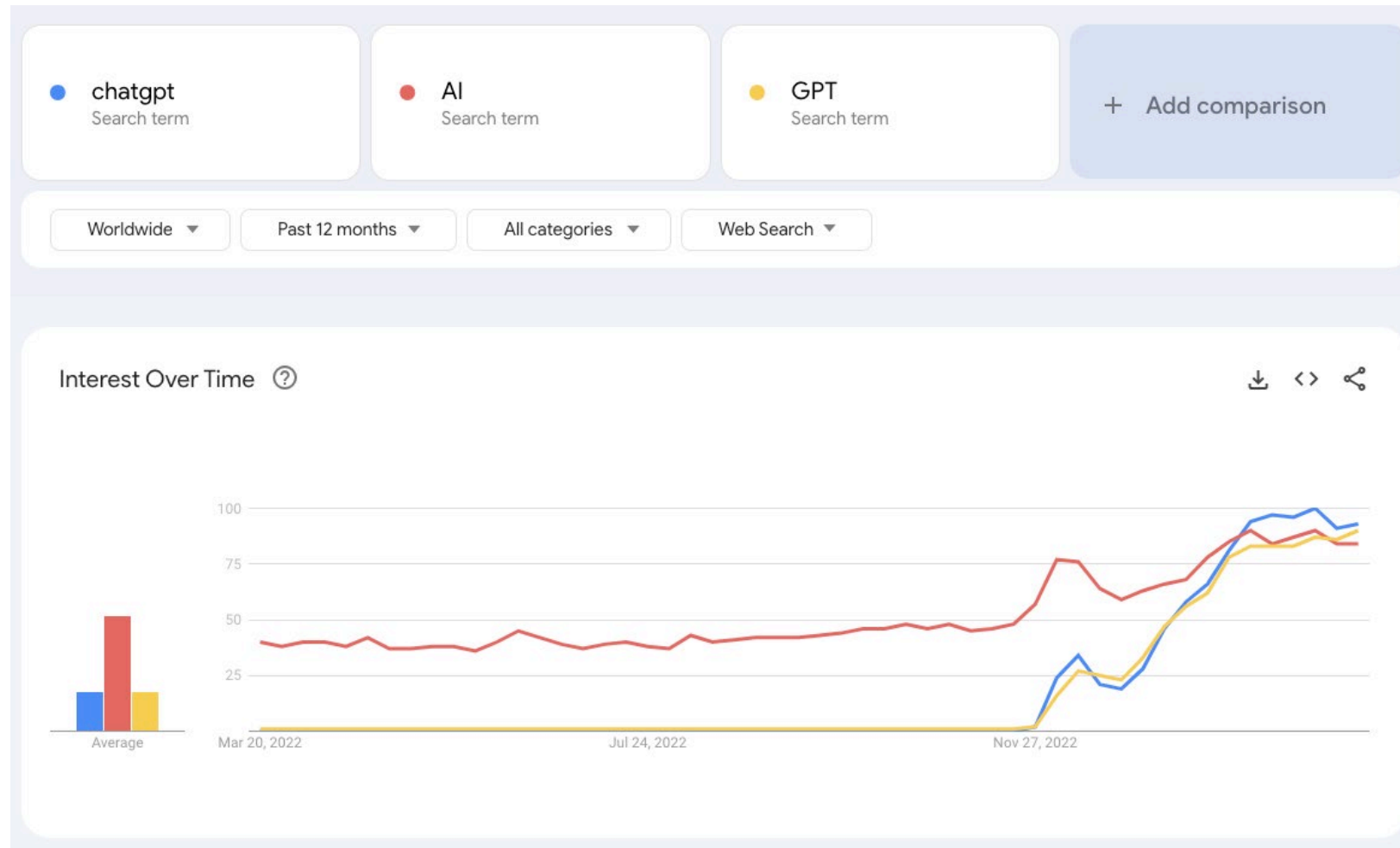
Outcomes are difficult to understand and test; data is large, non-integrated, and complex

...but AI capabilities are broad and can address a large array of challenges.

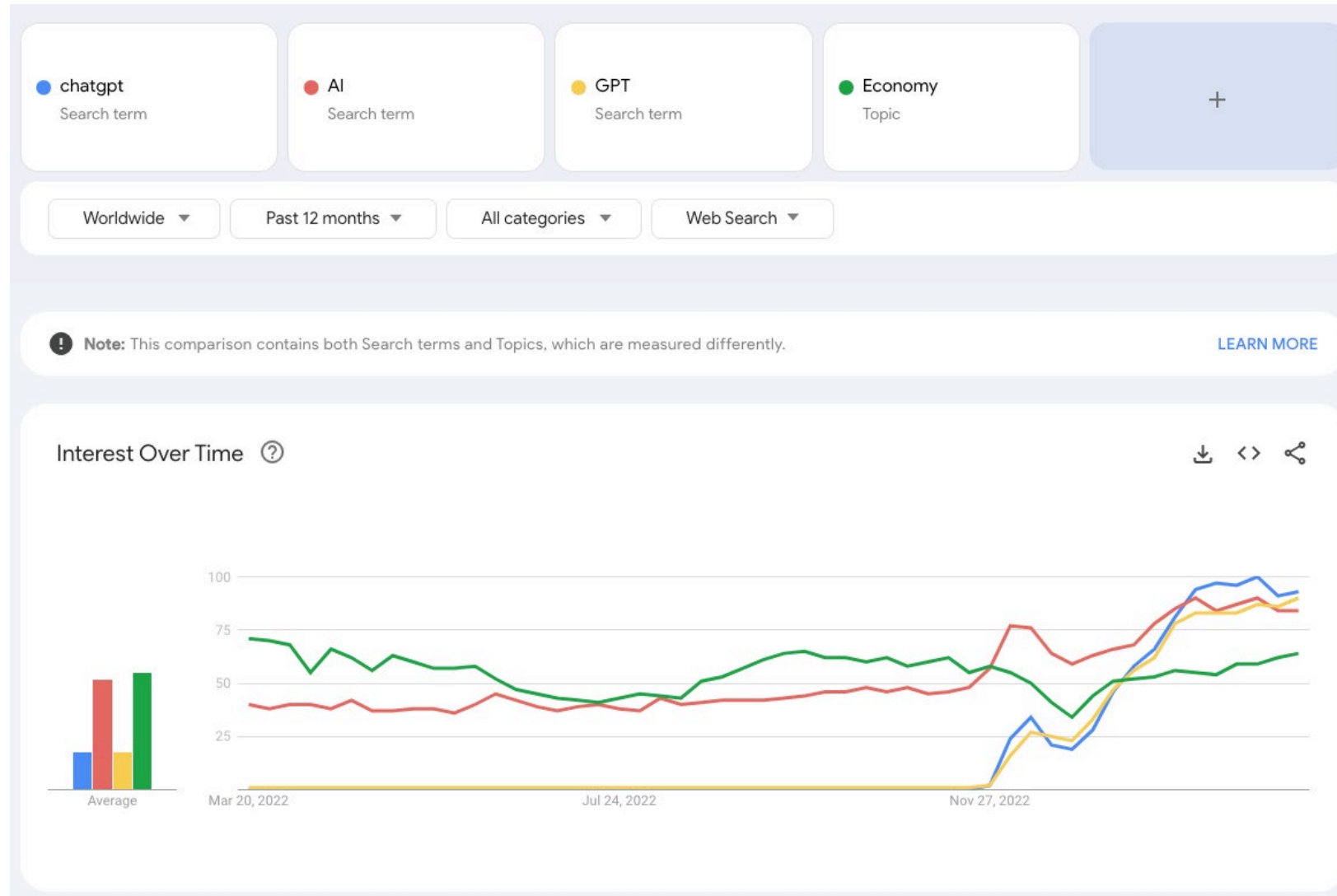
Deloitte's Classifications of Artificial Intelligence

	AUTOMATE – <i>Intelligent Automation</i>	<i>Having a bot execute processes that require little human judgement</i>	Gmail Autocomplete	
	SIMULATE – <i>Agent-Based Simulation</i>	<i>Using large volumes of data to create a visual model or representation of predicted outcomes in the form of a digital twin</i>	'Try On At Home' AR (e.g., L'Oréal's Style My Hair app.)	
	INTERACT – <i>Virtual Assistants</i>	<i>Machines engaging with humans in real-time, two-way dialog and learn human intent to provide advanced recommendations</i>	Smart Assistants (e.g., Amazon Echo)	
	PREDICT – <i>Predictive Analytics</i>	<i>Using large volumes of data to enable machines to generate a series of predicted or possible outcomes</i>	Movie recommendations	
	DETECT – <i>Computer Vision</i>	<i>Identifying objects and patterns through analysis of documents, pictures and video streams</i>	Airport Security using Eye Scanner	
	INTERPRET – <i>Natural Language & Speech Recognition</i>	<i>Training machines to read documents, convert text and speech to data and derive insights from data</i>	Auto-generated News Stories	
	CREATE – <i>Generative AI</i>	<i>Algorithms that can create new content, including audio, code, images, text, simulations, and videos</i>	Generated art (e.g., Lensa.ai)	

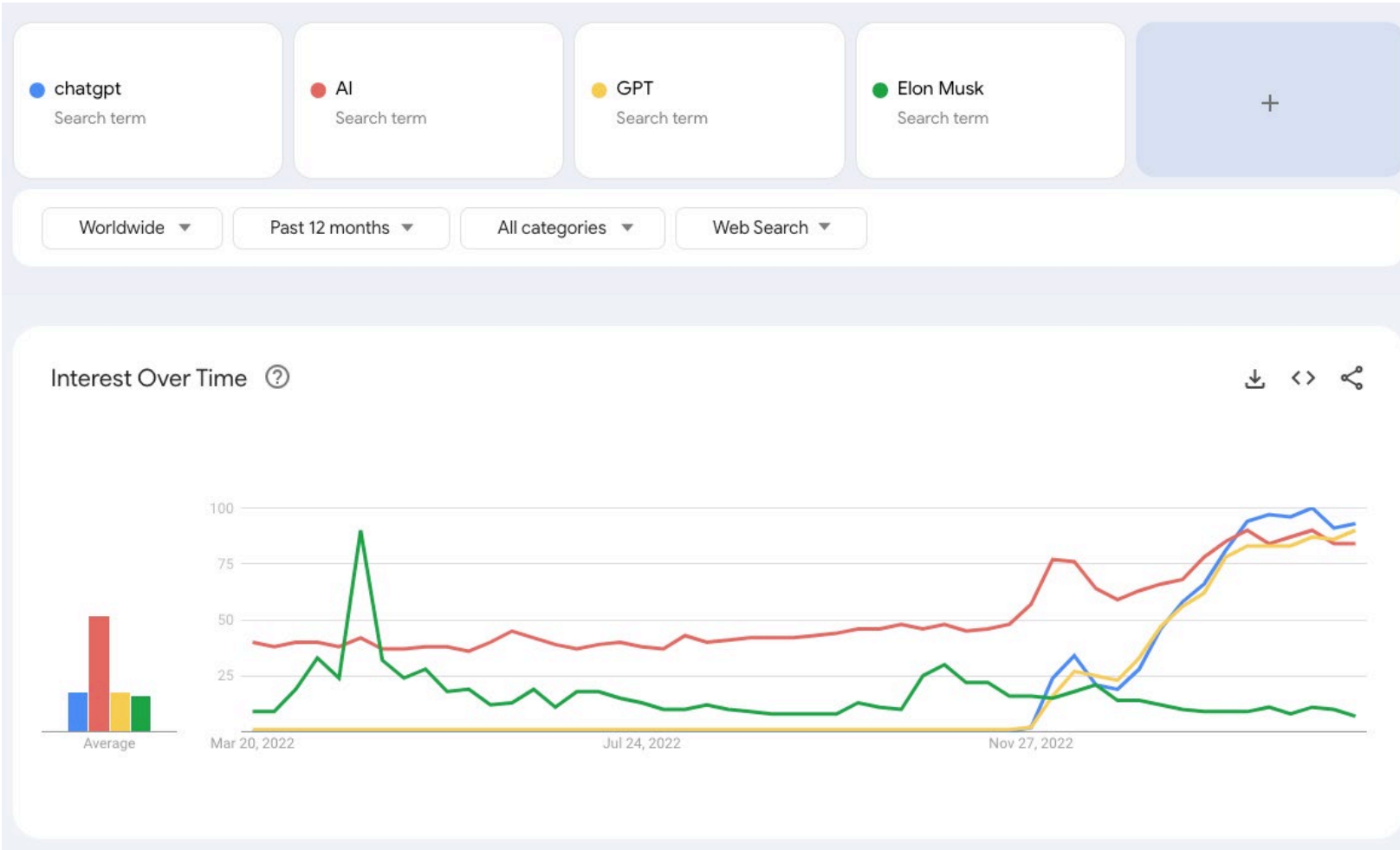
**ChatGPT & GPT
are now more
popular than AI
itself and has
doubled general
interest in AI as
well**



More popular
than anything to
do with the
Economy
combined

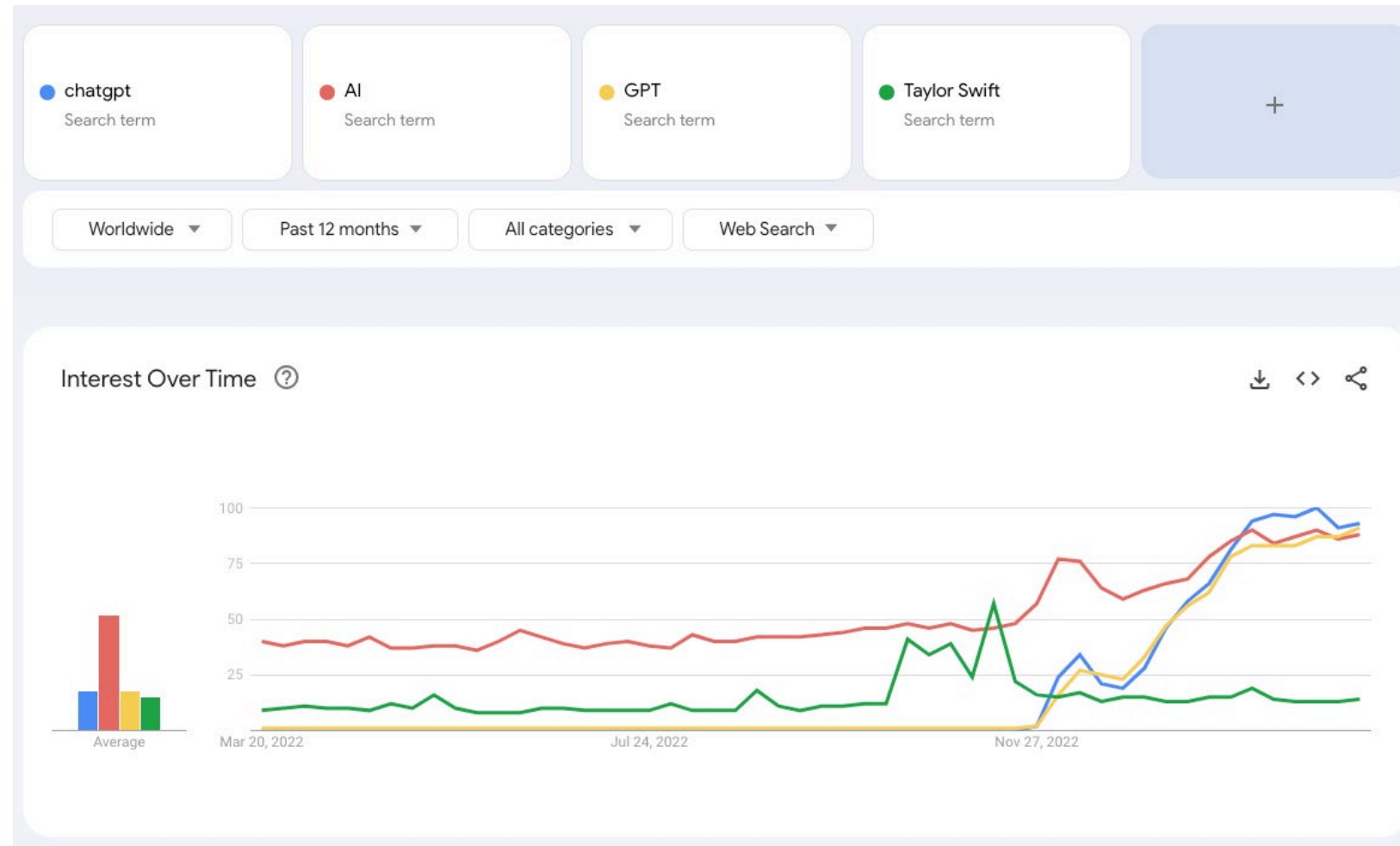


More popular than Elon Musk



**AND MOST
IMPORTANTLY...**

**More popular
than Taylor Swift**

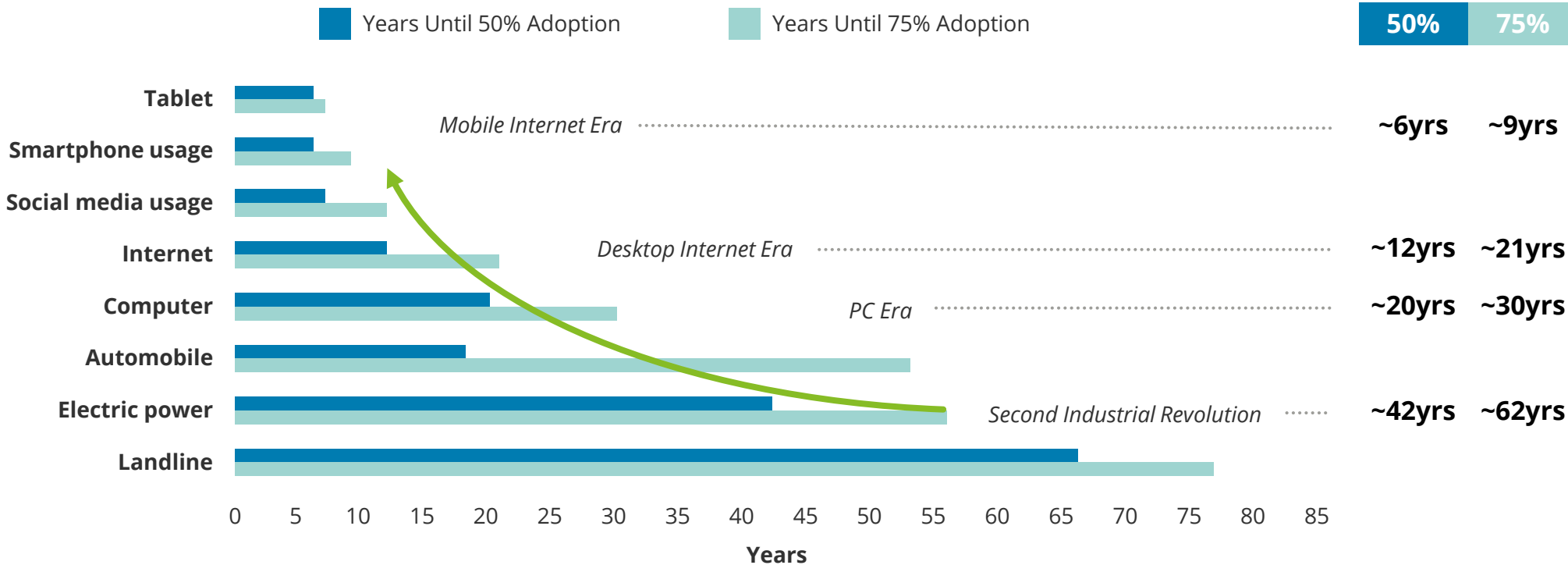


Generative AI has arrived and will be pervasive...

Faster Pace of Innovation

- Technology builds on itself, and increasingly, **new technologies are beginning to converge, speeding up the time to mass adoption**

Years Until 50% and 75% US Household Adoption of New Technology



Source: OurWorldinData, various websites, Morgan Stanley Research (Thematic Alpha: Data Era Spend – Generating Alpha and Productivity (22 Jul 2021))

Generative AI is the ultimate boardroom-level topic of the moment

MOVING AT UNPRECEDENTED PACE

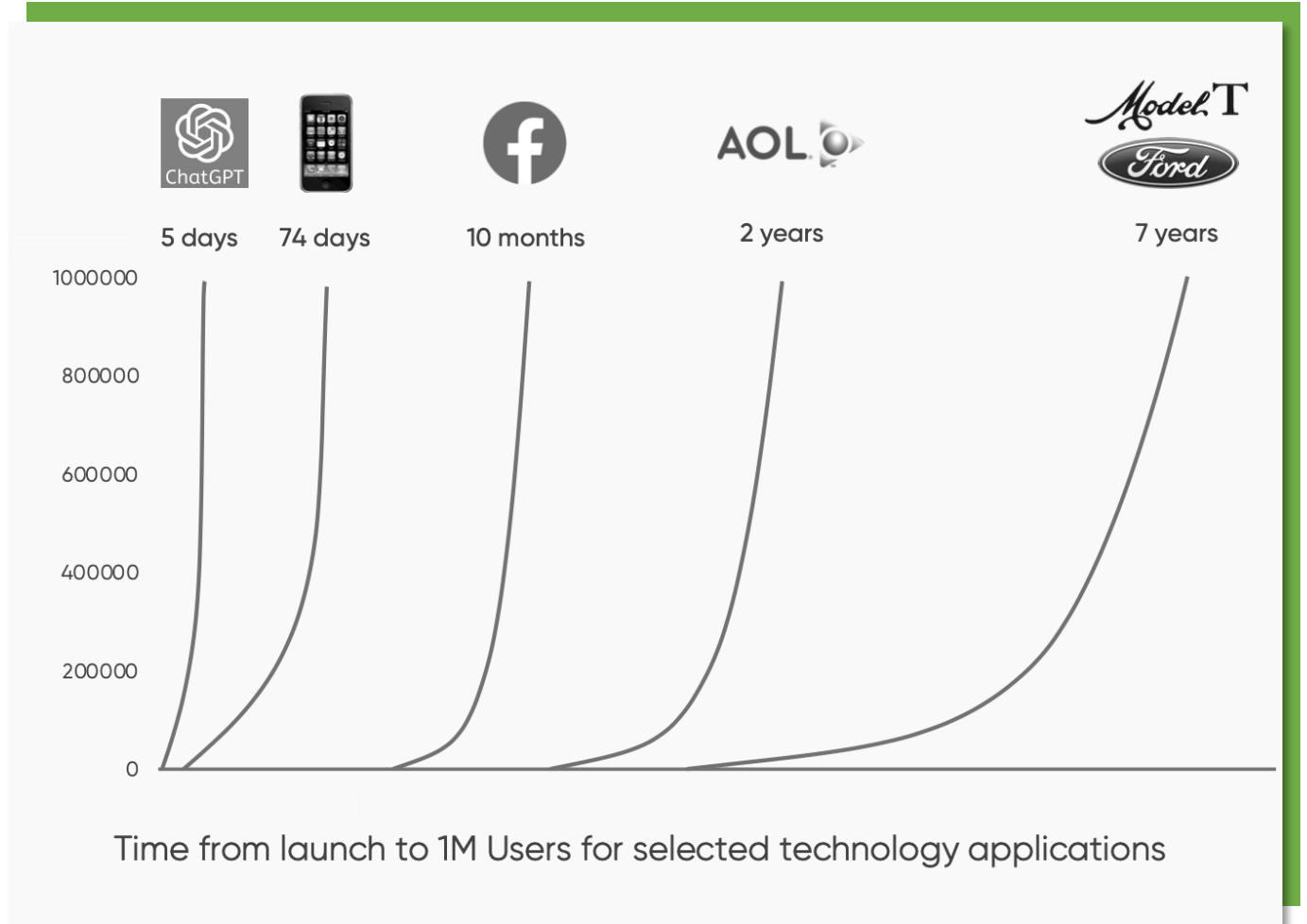
Gartner estimates that more than **10% of all data will be AI-generated** by as early as 2025, heralding a new age, the Age of With™.

ECONOMIC IMPACT

Generative AI expected to increase Global GDP by **\$7 trillion** (7% over 10-year period), with **1.5% growth in labor productivity**¹

MASSIVE MARKET

Analysts estimate the **market for Generative AI at \$200B by 2032**². This represents ~20% of total AI spend, up from ~5% today³.



How Do Generative AI Applications Work?

Imagine GPT as an infant...

... learning how to talk. The infant first learns basic patterns of how to talk from their parents, in this case, from the Internet.

Although it is generally very smart, ChatGPT, like infants, would output comments that contain misinformation, biases, or information that is not required. Users then give feedback to course correct. It is a **constant loop** of users giving **feedback**, as the GPT improves the quality of output.

Judgement and edits are crucial steps for humans, as we address the **alignment problem**: how can human beings create a model that reduces the number of biases and that aligns with our human values of what we'd want the world to look like?



TRAINING

ChatGPT learns from any information on the web including Wikipedia, web pages, etc. and excluding most recent data after 2021

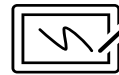


USER INPUT PROMPT

Just as you would in Google Search, type in what you'd want to know / do for you

E.g., "How do I make pancakes?" "Write a poem for valentine's day in the tone of Woody Allen"

GPT OUTPUT



GPT will answer your question / request

E.g., "First you put 100g of flour..." "Roses are red...etc"



USER FEEDBACK

Specify or modify the prompt or give a thumbs up / down feedback to rate the output.

E.g. "How do I make pancakes for a calorie conscious mom with no eggs in the fridge?"

GPT UPDATED OUTPUT



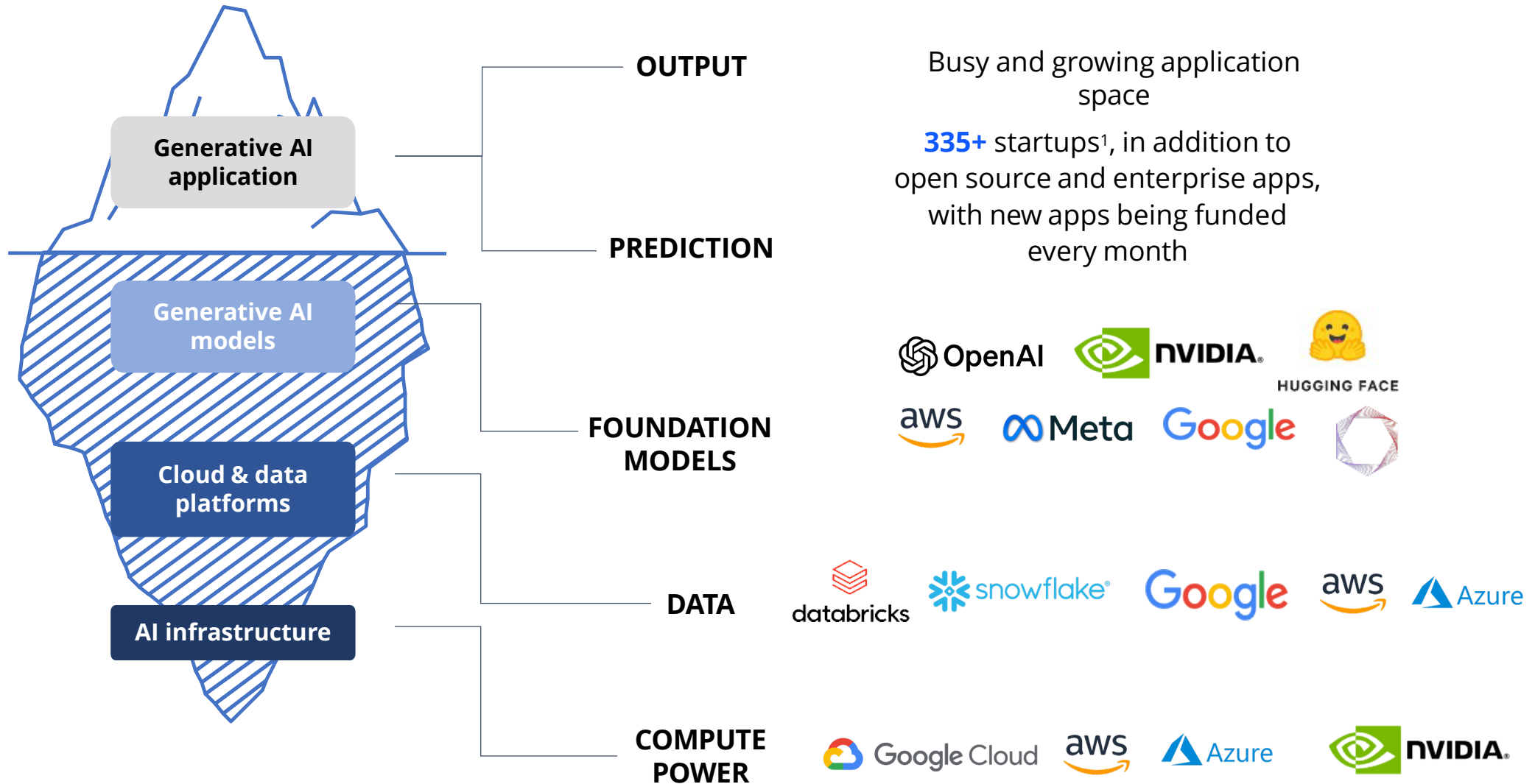
GPT will provide you with, hopefully, an answer closer to what you want



USER EDIT/JUDGEMENT

Make the judgement on the quality of content (text, image etc.) and modify the prompt again if/as needed

Gen AI is built on a robust tech stack; the landscape is evolving quickly and its too early to try picking winners and losers



¹ Source: CB Insights

A solid blue vertical bar is located on the far left side of the image, extending from the top to the bottom.

Harnessing the Power!



...and to drive “Productivity” in the Enterprise

Sales, Marketing, Customer Service

- Sales and Marketing Content Generation
- Content Localization, A/B Testing
- Personalization
- Product Discovery, Sales & CPQ assistance
- Voice and Chat Assistant

IT, Coding, Cyber, Service Desk

- Autonomous Coding (generation, testing, migration, explanation)
- Data management
- Incident management



Human Resources

- Learning & Development
- Talent Acquisition
- Employee support systems/Policy knowledge management

Finance

- Investor Relations
- Financial Reporting and Analysis
- Ad hoc querying
- Policy Knowledge Management

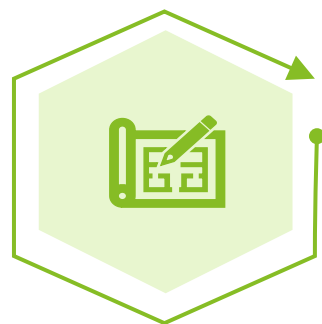
+30% Improvement in **Software Project Velocity** from GenAI assisted coding

~30%

Reduction in **customer service costs** from advanced Voice Assistants (est.)

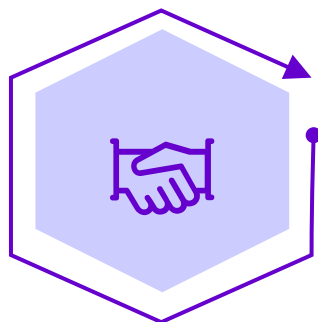
50-90% Reduction in effort **gathering & analyzing data** from knowledge mgmt. systems (est.)

A methodical approach is necessary to drive strong returns out of AI investments



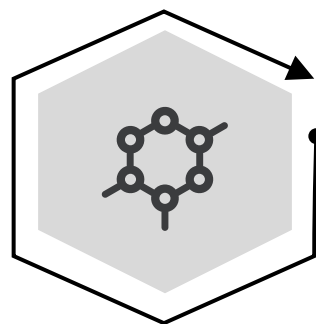
Strategy & Prioritization

How do we prioritize and plan to integrate Generative AI use cases & capabilities into our enterprise business strategy?



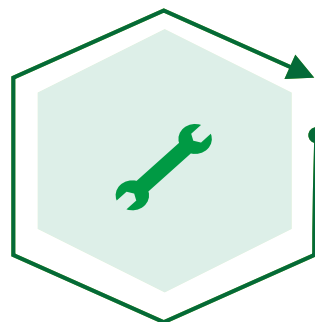
Partnerships & Alliances

Who are your partners and how will you orchestrate the growing partner ecosystem?



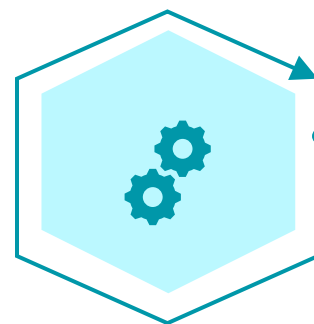
Platform, Data & Integration

How do we integrate LLMs into our enterprise business and data architecture?



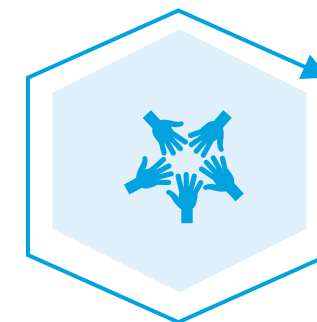
Use Case - Model Selection

Which models to use for which use case, should we build or buy the LLM?



Operating Model

What organizational structures, capabilities, skillsets and processes do we need to support Gen AI?



Trustworthy AI

How do we ensure that Gen AI usage is secure, transparent and responsible?



Legal

What are emerging regulatory concerns around Generative AI?

Capabilities Automation

Generative AI

Business Insights

AI Engines

App Library

ENTERPRISE APPS

ReportGen

CUSTOM WORKFLOWS

QUERY DOCUMENTS

BI Reporting & User Feedback

Deloitte AI Cloud

AI Pipeline Builder

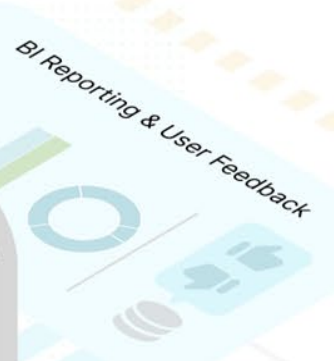
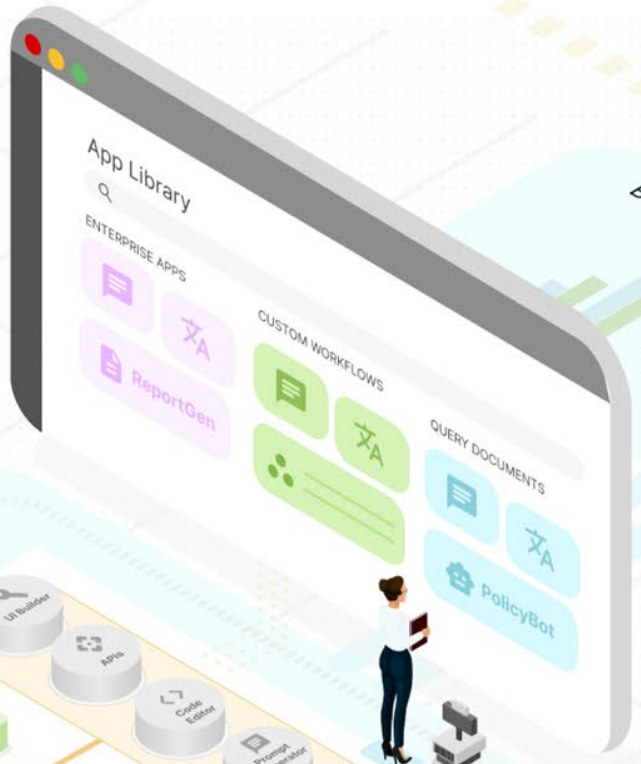
Access Management

Benefits

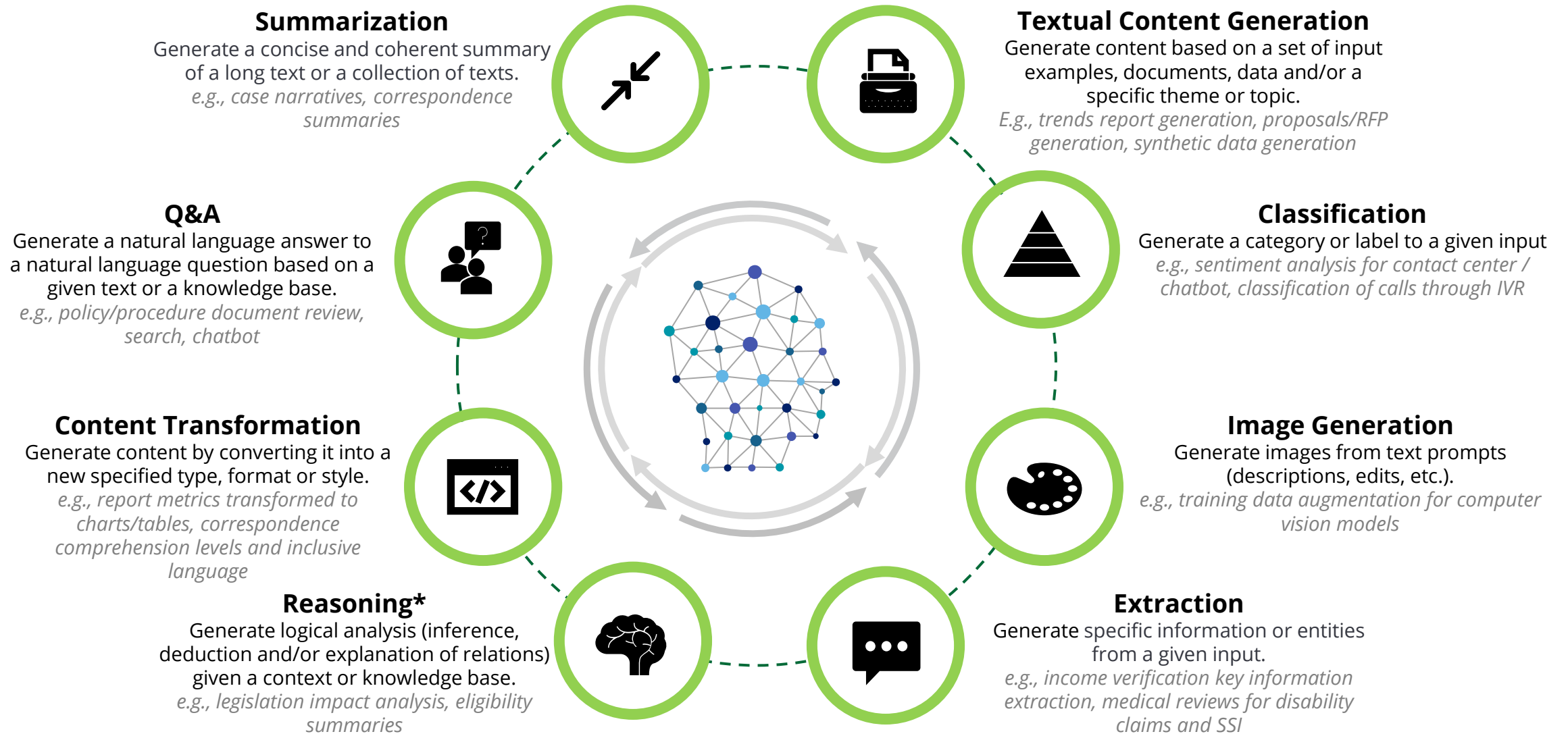
Secure

Accelerate

Save



Generative AI has a wide range of uses, and others are yet unimagined



* Current Generative AI reasoning capability remains to be tested and expected to improve for future models

State Government Focus for Generative AI Use Cases



COMMUNICATIONS & REPORTING

- Document summarization and categorization
- Chatbot / virtual assistant dialogue generation
- Strategic communication generation
- Policy and regulation change analysis and summary
- Report generation and impact analysis
- Video or image editing and generation



HUMAN RESOURCES

- Personal onboarding assistant
- Compensation analysis
- Workforce skill analysis
- Resume analysis
- Metaverse 3D workforce experience
- Metaverse 3D workforce upskilling



CONTRACTING & PROCUREMENT

- Language translation
- Supplier/Provider identification and evaluation
- Automated bidding and proposal evaluation
- Fraud detection and prevention
- FAR adherence and recommendations
- Contract adherence & anomaly detection



GOVERNANCE & OPERATIONS

- Intranet search (knowledge management)
- Process analysis
- Training for new team members
- Document inventory analysis
- News and media summaries
- Sentiment analysis for workforce



INFORMATION TECHNOLOGY

- Legacy code summarization & translation
- Peer review for optimized code writing
- Training on new technologies
- Test automation and test scenario creation
- Development lifecycle documentation
- Code generation across languages/frameworks/CSPs



FINANCE & ACCOUNTING

- Invoice processing and payment automation
- Risk management
- Forecasts and planning
- Financial report analysis
- Regulation and oversight analysis
- Fraud, waste, and abuse prevention

Making an Impact in Health and Human Services



**Julia
(Resident)**

- Feels extremely stuck and overwhelmed
- She lost her job last week and is the primary provider for her family
- She has a fifteen-year-old granddaughter and a twelve-year-old grandson she cares for
- She has heard of food stamps, but has no idea where to start

How Does AI Help?

- Julia learns that she’s not only eligible (potentially) for SNAP, but she should also apply for Medicaid, TANF, LIHEAP, and HAF
- She fills out a simple application to officially apply for all the above benefits
- She’s received a tailored text message letting her know the status of her application
- Benefits are approved and received within a few days helping to alleviate her concern
- Julia is notified of employment and training services

Considerations for AI Success

- Assess and establish standards for safeguarding PII data
- Prompt engineering/testing to improve accuracy, reliability, and bias reduction
- Establishing guardrails to connect residents with the appropriate resources and services
- Alignment of eligibility rules and criteria based on recommended program (Medicaid, SNAP, TANF)

Use Cases

Connect clients to resources and discovery services available

Provide text, email, and online notices and alerts

Provide multi-language text generation support for translating notices

Enable clients with question-answering capability against websites and State data

Develop custom checklists that help customers maintain their benefits

Assist caseworkers, business analysts, and executives answer complex policy questions

Making an Impact in Workforce Development



Robert
(Workforce Supervisor)

- Never feels his work is done
- Feels unorganized with the endless influx of applications, tasks and cases
- Extremely frustrated with trying to learn new policies and regs
- Robert’s staff finds communication with clients to be challenging due to language barriers

How Does AI Help?

- Robert uses the new Policy engine tool to research if refugees who are new residents of his state are eligible for workforce support
- Robert’s staff can now access all their tasks and cases in a single integrated workload dashboard
- He’s able to improve case coordination for staff to ensure their caseloads are manageable
- The new multilingual translation services helps his staff communicate with multilingual jobseekers

Considerations for AI Success

- Review existing policy documentation to confirm accuracy/ reliability
- Establish data transparency to clearly identify the reference/ source documentation for the policy answer
- Separate LLM security protocols are required for policy data (public) vs. case copilot data (private PII)
- Implement continuous monitoring processes to track, train, and tune the workforce solutions
- Test translation services using prompt validation or Human in the Loop AI testing

Use Cases

Assist workers by interpreting policy or system questions	Provide tasks and caseload information summarized in a single view	Interpret policy, systems, and ops manuals to answer worker questions
offer industry-specific questions and facilitates a training exercise simulating a live job interview	Provide personalized guidance by using information from existing policy/procedure documents	Answer FAQs related to tickets, resolving help desk issues, and summarizing help desk calls.

Making an Impact in Higher Education



**Janelle
(Student)**

- First in her family to attend college, managed her FAFSA independently
- She struggled to juggle numerous grants and loan applications
- Extremely frustrated by the need to visit multiple offices for financial assessments and scholarship applications
- Constantly anxious about the impending debt

How Does AI Help?

- Janelle uses the new scholarship matchmaker to identify new scholarship opportunities applicable to her
- Tuition Payment Simulation allows Janelle to estimate her tuition costs compared to other universities
- The student job connector relieves some of the debt stress from Janelle, allowing her to find on/off-campus jobs to begin paying for college

Considerations for AI Success

- Identify and investigate potential AI-induced hallucination risks
- Promote and ensure ethical use of AI technologies in decision-making processes
- Enhance measures for maintaining confidentiality and privacy in AI-driven systems when dealing with FAFSA related data
- Prompt engineering/testing to improve accuracy, reliability, and bias reduction

Use Cases

AI-powered scholarship platform that matches student profiles to scholarship opportunities

Simulate tuition payment options to showcase options for students to fund their education and decide between various institutions, and options for schools to receive tuition money promptly.

Leverage AI to help students identify and match with various on-campus and off-campus jobs and paid opportunities.

Making an Impact in Transportation



**Anthony
(IT Operations)**

- Data wiz and has been writing reports and queries for the entire transportation department
- Concerned that when he gets promoted no one will be able to understand their data
- Data quality and data management continue to be ongoing challenges
- Interested in AI but not sure where to start

How Does AI Help?

- Anthony uses our Discover AI solution to quickly generate SQL based on text
- The Report Generation and Data Summarizer tools enable him to analyze and build reports on unstructured and semi structured data
- He can document and create a repository of common queries with the AI Catalogue
- Anthony can automatically obfuscate/ mask data in lower environments and create synthetic data for testing with Nudge Engine
- He attends an in-person technical GenAI training – specifically how to use Open AI within Azure

Considerations for AI Success

- Design and implement a secure architecture for AI systems to mitigate vulnerabilities
- Assess and establish standards for safeguarding sensitive data
- Address bias in AI algorithms to ensure fairness and equity
- Establishing guardrails to detect and prevent malicious activities in AI systems

Use Cases

Assist technical staff in generating SQL queries/ syntax based on English language questions

Identify actions within a system and nudges user to perform actions accordingly

Automatically generate reports and dashboards based on data

Bring together large data sets/documents and summarizes key information

Provide the ability to catalog and store both queries and models in single platform

Use natural language processing to answer policy and case questions.

Making an Impact in Tax / Revenue Services



Emily
(Audit Supervisor)

- Tax analyst that worked her way up to audit supervisor
- Concerned that her team is missing the right cases to audit
- Agency has had challenges with data management
- Has heard about AI but unsure of how it's different from GenAI and how this technology could be used

How Does AI Help?

- The Intelligent Case Management supports Emily and her staff with creating a prioritized list of tax files that should be audited based on the actionable details that are identified
- When taxpayers receive correspondence indicating that they have been selected for a tax audit, the Intelligent Audit Engagement virtual support chatbot identifies what information and documents are being requested
- Emily uses the Intelligent Back Office to improve workload management and ensure staff are focused on complex audit cases and providing audit support to ensure taxpayer compliance

Considerations for AI Success

- Assess and establish standards for safeguarding PII data
- Prompt engineering/testing to improve accuracy, reliability, and bias reduction
- Continuously update and improve the machine learning models to adapt to changing tax laws, regulations, and patterns

Use Cases

Use GenAI to parse tax case files to identify actionable details such as excessive deductions, unreported income, excessive expenses, and missing income, which are then sorted and prioritized to identify the most likely cases that should be audited

Frontline virtual support chatbot that helps taxpayers who have been selected for an audit understand the scope of the audit, what information and documents are being requested, and what information and documentation should be included in the audit response

Use GenAI to augment the workforce to complete previously tedious back-office functions related to decisions, HR, procurement, and logistics to allow for staff to focus time and attention on audit support to ensure taxpayer compliance

Bringing AI to Life through Real World Applications

AI and GenAI are helping to **reimagine and enhance** the **experience** for customers, staff and supervisors, students, and technology staff.

Health & Human Services



**Julia
(Resident)**

- Julia learns that she's not only eligible (potentially) for SNAP, but she should also apply for Medicaid, TANF, LIHEAP, and HAF
- She fills out a simple application to officially apply for all the above benefits
- She's received a tailored text message letting her know the status of her application
- Benefits are approved and received within a few days helping to alleviate her concern
- Julia is notified of employment and training services

Workforce Development



**Robert
(Workforce Supervisor)**

- Robert uses the new Policy tool to research if refugees who are new residents of his state are eligible workforce support
- Robert's staff can now access all their tasks and cases in a single integrated workload dashboard
- He's able to improve case coordination for staff to ensure their caseloads are connected through data
- The new multilingual translation services helps his staff communicate with multilingual jobseekers

Higher Education



**Janelle
(Student)**

- Janelle uses the new scholarship matchmaker to identify new scholarship opportunities applicable to her
- Tuition Payment Simulation allows Janelle to estimate her tuition costs compared to other universities
- The student job connector relieves some of the debt stress from Janelle, allowing her to find on/off-campus jobs to begin paying for college

Transportation



**Anthony
(IT Operations)**

- Anthony uses GenAI solution to quickly generate SQL based on text
- The Report Generation and Data Summarizing tools enable him to analyze and build reports on unstructured and semi structured data
- He can document and create a repository of common queries with an AI Catalogue
- Anthony can automatically obfuscate/ mask data in lower environments and create synthetic data for testing with recommendation type capabilities

Tax/Revenue



**Emily
(Audit Supervisor)**

- Using AI in Case Management supports Emily with creating a prioritized list of tax files that should be audited
- When taxpayers receive correspondence indicating that they have been selected for a tax audit, the an AI driven audit engagement chatbot identifies what information and documents are being requested
- Emily uses the AI in the backoffice Back Office to improve workload management and ensure staff are focused on complex audit cases and providing support to ensure compliance

5 lessons learned from year 1 of working with 70+ organizations on their GenAI journey

1

Understand what Generative AI is, and what it isn't

Generative AI is only one type of Artificial Intelligence, and learning the fundamentals helps calibrate expectations

2

Prioritize data quality & data rights to drive impact

The quality of data ingested by Gen AI models is directly correlated to the model's results

3

Consider a Human + Machine approach to Generative AI

Cultivate a culture of trust, both for the outputs of AI models, and within the workplace, as you introduce Gen AI as a new coworker

4

Embrace the uncertainty

The landscape is changing quickly and technology is evolving quickly. Instead of looking for certainty, it's critical to embrace uncertainty, experiment, and to learn and buy options on that transformative future

5

Don't wait to get started

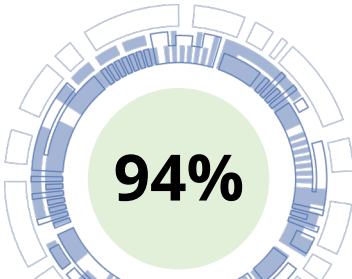
Organizations have seen success in standing up a Generative AI COE, which can oversee all components of a Generative AI, from strategy through execution

Managing Risk and Ethical Considerations in AI

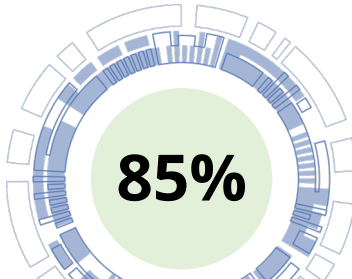
AI Is Driving Broad Evolution in Government, Society, and Business



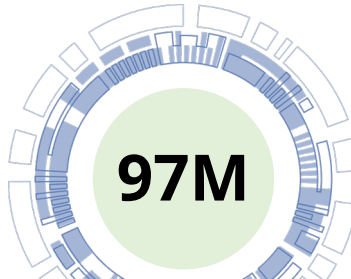
Deloitte’s 5th annual “State of AI in the Enterprise”¹ report reveals that global leaders agree AI is essential to their organization’s success more now than ever before



indicate that **AI will be critical to the success of their organization** over the next five years



expect increased investment in AI tech and talent over the next year



jobs will be created across 26 countries by 2025

¹ Based on Deloitte’s [State of AI in the Enterprise](#), 5th Edition – survey of ~2,620 leaders

State AI Policy Landscape

BACKGROUND



EDUCATION

Policymakers at state executive and legislative levels want to **understand how AI is being used in their own states** and educate themselves before legislating on the topic.



LEGISLATION

During the 2023 session, **artificial intelligence legislation** was introduced in **over 22 states and territories**.



WORKING GROUPS

Legislation enacted and executive orders issued in 2023 have pertained to creating **working groups and committees tasked with studying AI and producing policy recommendations**.

Example State Actions



California: Executive order issued to ensure deployment of ethical and responsible GenAI within state agencies.



Connecticut: Legislation established a working group to develop best practices for ethical and equitable use of AI in state government; required an inventory of state systems that employ AI.



Maine: MaineIT established a moratorium for at least 6 months on GenAI in order to conduct a risk assessment.



New Jersey: Executive order issued that establishes AI Task Force to study emerging AI technologies, to issue findings on their impacts, and offer recommendations on ethical use of them.



Oklahoma: Executive order issued to create the Governor's Task Force on Emerging Technologies to develop policy recommendations on responsible deployment of AI and GenAI.



Pennsylvania: Executive order issued to establish responsive and responsible standards for governance of Gen AI when used by state agencies.



Texas: Legislation established advisory council to study and monitor AI systems developed, employed, or procured by state agencies.



Virginia: Executive directive issued to ensure the responsible, ethical and transparent use of AI technology by state government.



Generative AI Market Pop Quiz!

Get ready to test your knowledge on some quick questions on GenAI!

Question #1: Which picture was actually generated by AI?

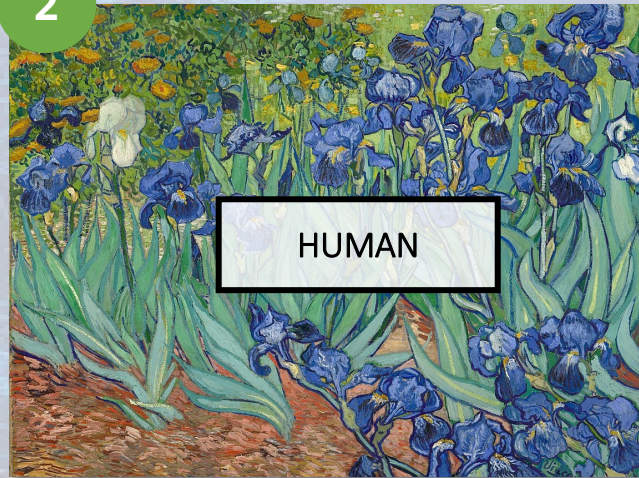


Question #1: Which picture was actually generated by AI?

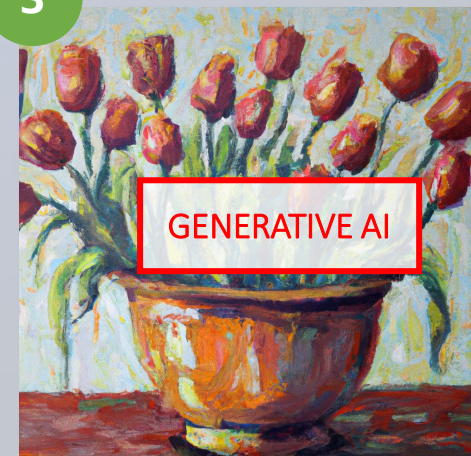
1



2



3



Question #2: What **part of this image** was generated by AI (e.g., trees, mountain, etc.)?



Question #2: What **part of this image** was generated by AI (e.g., trees, mountain, etc.)?



Question #3: Which of the **following pictures** were created through generative AI?

1



2



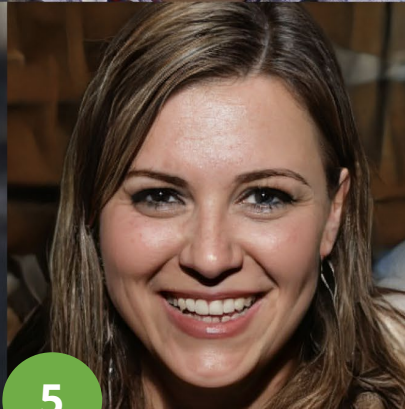
3



4



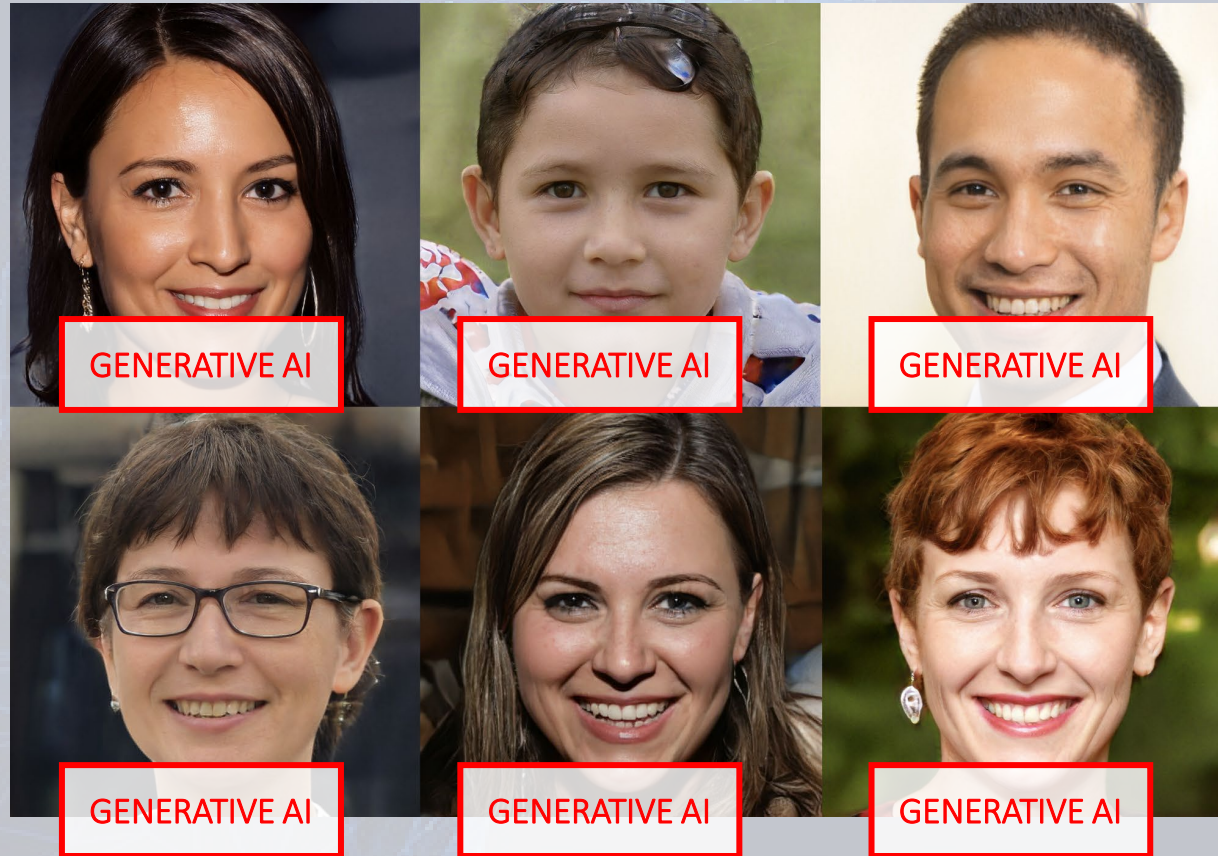
5



6



Question #3: Which of the **following pictures** were created through generative AI?



Ethical, Functional, Legal and Security Risks



Bias

Bias in; bias out. If the training data is biased (e.g., over/under-representation of a population cohort, sexism, racism), then the outputs generated could also exhibit biases. Bias reductions in the training data and/or human supervision during model training is needed.



Text Formatting

Models are good at understanding text, but they struggle when the data are in irregular formats or when the position of the text on the page (e.g., infographic, presentation slide) is relevant to the context and understanding. Other emphasis generators, such as bolded text, font color, etc., don't play a role yet.



Hallucination

Models might output statements that are factually false. Sources and citations are unavailable for most models. Users should be conscious that outputs could be inaccurate and perform due diligence to validate generated content.



IP Protection & Infringement

SaaS-AI organizations may use prompt payloads to train future versions of the base model, potentially including confidential data that could expose the user to IP infringement claims – how could this affect your organization's competitiveness in the market?



Secure Infrastructure

External generative AI tools such as ChatGPT introduce an insecure extension of the environment, such as the inability to determine and how that suitable controls are in place for regulatory and policy compliance.



Ethical Use

Is the AI being used in a manner consistent with the purpose of the overall exercise? Is a human being brought into the loop to decide whether the AI's suggestion needs adjustment before actual use or whether the use of AI is ethical (e.g., submitting an AI-generated essay as your own)?



Model Performance

Foundation Models are comprised of billions of parameters (model size) and trained on petabytes of data. In theory, the larger the model, the better the output. Foundation Models take time to produce outputs, which may limit real-time use cases.



Cost

Foundation models generally offer a pay-as-you-go billing mechanism, and the cost per use of sophisticated models is materially significant. Fine-tuning the biggest model and running large documents through several times could quickly run up a bill of tens of US \$1,000s.



Malicious Behavior

To maintain operations and customer trust, proactively minimizing risk from malicious behavior on the network is critical. For example, a customer service bot revealing confidential information to a hacker either by prompt or unintentionally.

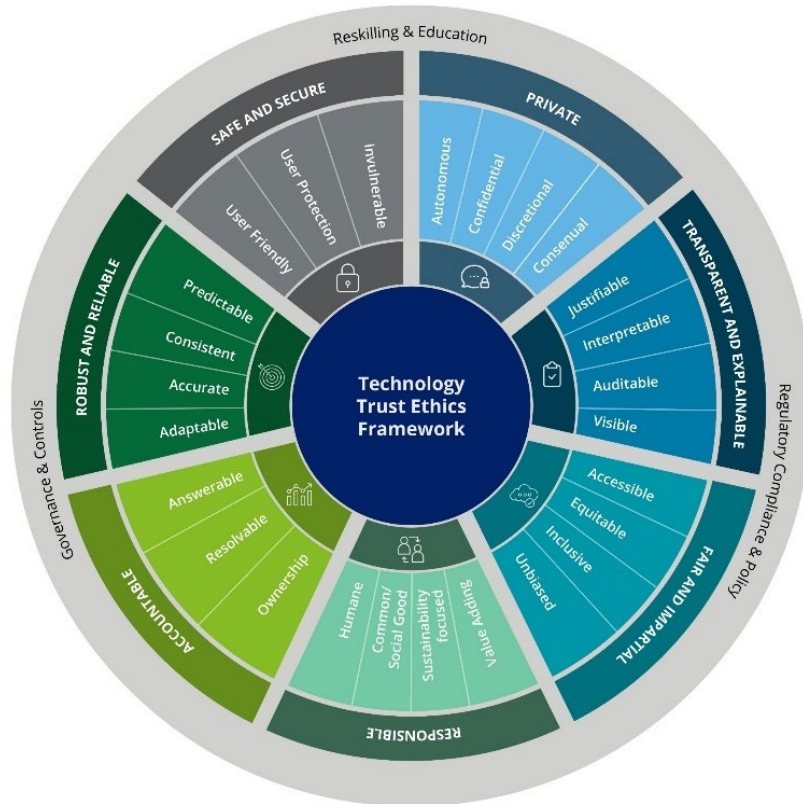


Confidentiality & Privacy

Generative AI Models are built on data sharing which makes it challenging to maintain an individual's privacy rights. Consent for data used (confidential information, personally identifiable information) is necessary, but poses concerns around an individual's right to consent or be forgotten.

Technology Trust Ethics (TTE) Framework for Risks & Limitations Associated with Generative AI

TTE framework helps organizations develop **ethical safeguards** across 7 key dimensions to build, deploy and commercialize generative AI applications—a crucial step in managing the risks and capitalizing on the returns associated with generative AI. Our framework helps to ensure the technology is:



The framework can be used for all technologies including AI, data, Non-Fungible Token (NFT), Metaverse, 5G, Quantum computing etc.

Private User privacy is respected, and data is not used or stored beyond its intended and stated use and duration; users are able to opt-in / out of sharing their data.

Transparent & Explainable Users understand how technology is being leveraged, particularly in making decisions; these decisions are easy to understand, auditable, and open to inspection.

Responsible The technology is created and operated in a socially responsible manner.

Fair & Impartial The technology is designed and operated inclusively in an aim for equitable application, access, and outcomes.

Safe & Secure The technology is protected from risks that may cause individual and / or collective physical, emotional, environmental, and/or digital harm.

Robust & Reliable The technology produces consistent and accurate outputs, withstands errors, and recovers quickly from unforeseen disruptions and misuse.

Accountable Policies are in place to determine who is responsible for the decisions made or derived with the use of technology.

Approaches to Address Risks & Limitations

While not exhaustive, below are 4 key steps to take in response to generative AI risks and limitations

1

AI Literacy

Host informative sessions that focus on educating business and tech leaders, as well as data users, about the responsible utilization of generative AI. These sessions aim to enhance their understanding and knowledge of ethical practices in generative AI usage.

2

Continuously evaluate the changing regulatory landscape

Stay updated on the evolving regulatory requirements related to generative AI and ensure effective communication and integration of those requirements into policies, processes, and implementation approaches.

3

Establishing appropriate audit process for vendors

Implement an appropriate level of testing and validation, whether independent or internal, to allow for visibility and transparency of the impact of risks when generative AI models are implemented.

4

AI Governance Framework

Establish an operating model where experts can help and guide the organization to use generative AI solutions appropriately by accessing the implications for existing processes, proactive policy enhancements, and incorporating necessary safeguards and guidelines into those processes.

Big AI promises are being made but some expectations may be unrealistic



Questions?

