Gartner Webinars

Gartner delivers actionable, objective insight, guidance and tools to enable stronger performance on your organization's mission critical priorities



Enhance your webinar experience







Ask a Question

Download Attachments

Watch Again





How CIOs Can Calculate Business Value and Cost for Generative Al Use Cases



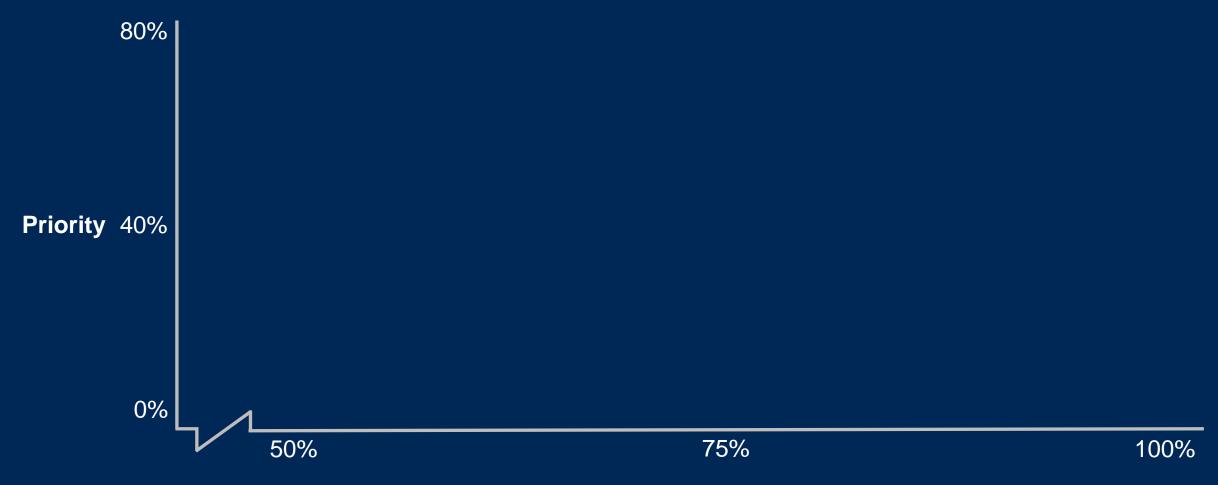
Rita Sallam

Distinguished VP Analyst



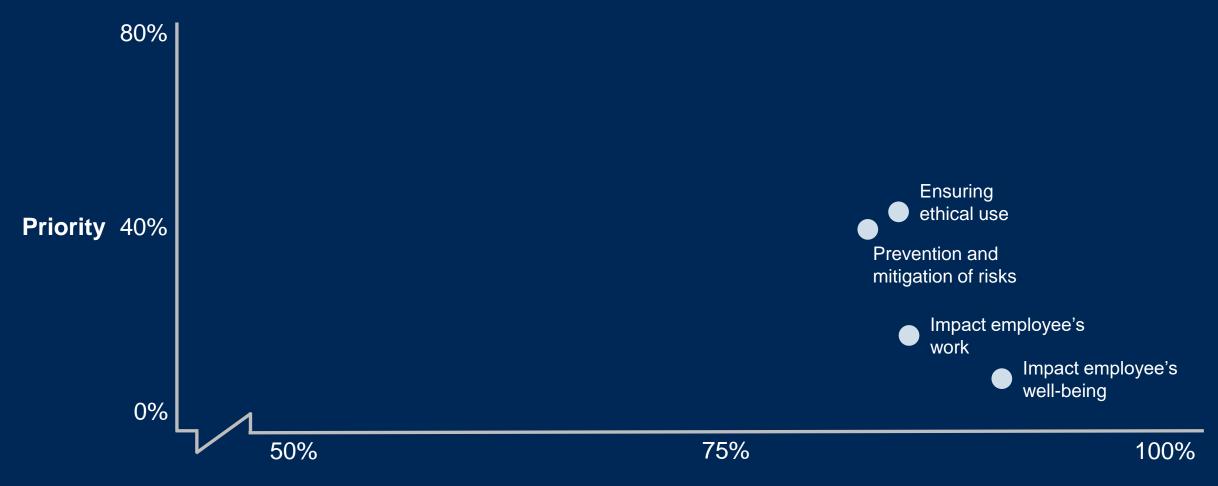


How Are You Performing Against Expectations?



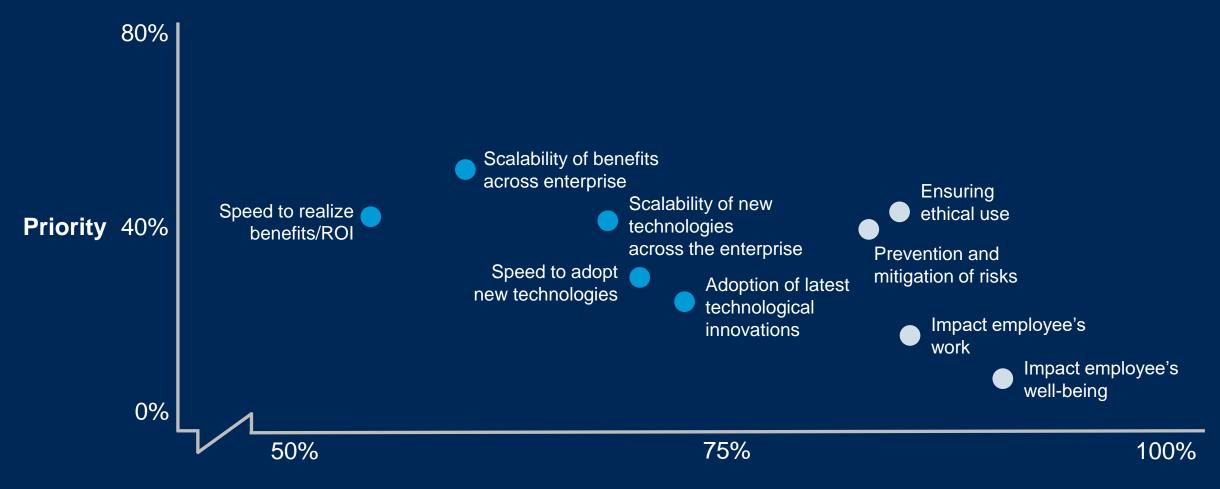


Some Things Are Going Well



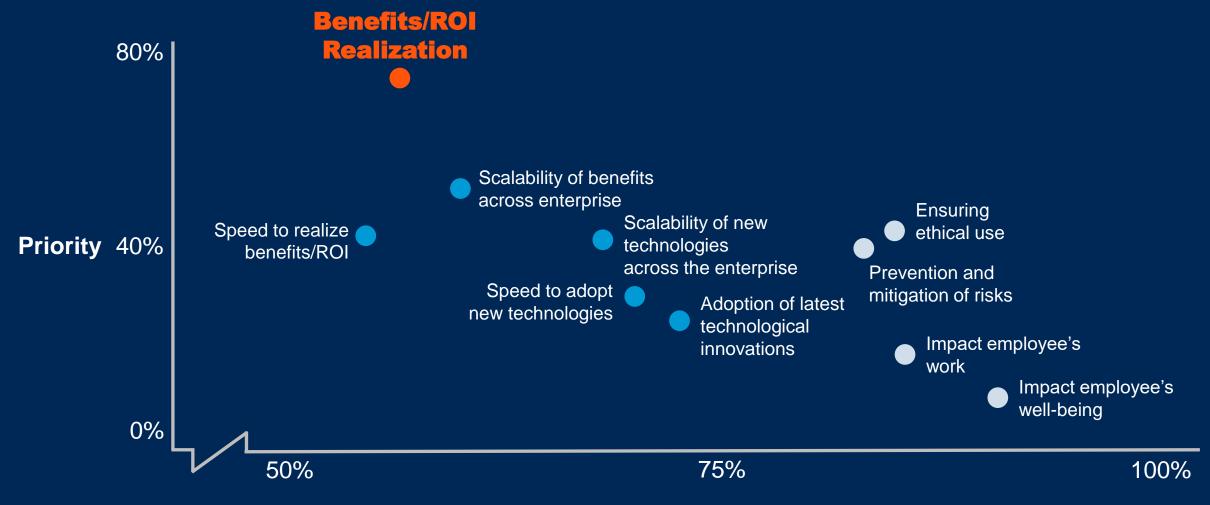


Others Are OK



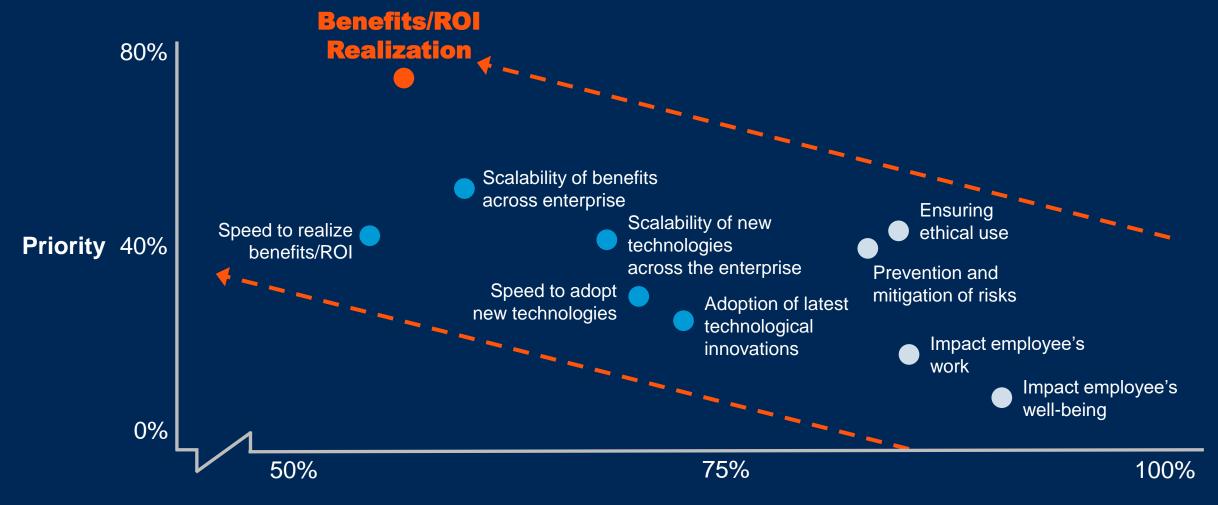


ROI Is Most Important, but Most Challenging





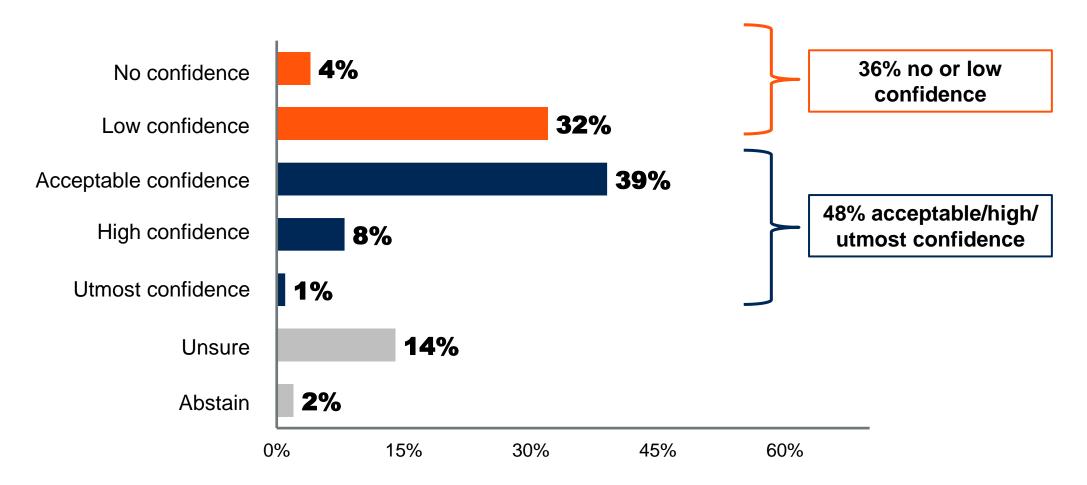
ROI Is Most Important, but Most Challenging





Levels of Confidence: Calculating Costs for AI in the Enterprise

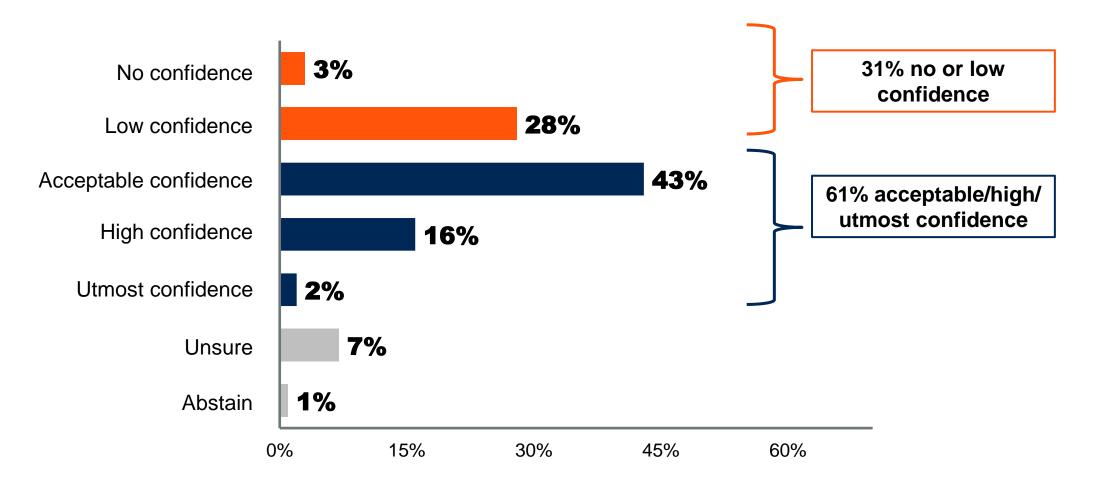
Q: What is the level of confidence you have in your organization's approach to **calculating cost** for Al initiatives (use cases)? Percentage of respondents





Levels of Confidence: Assessing Risk for AI in the Enterprise

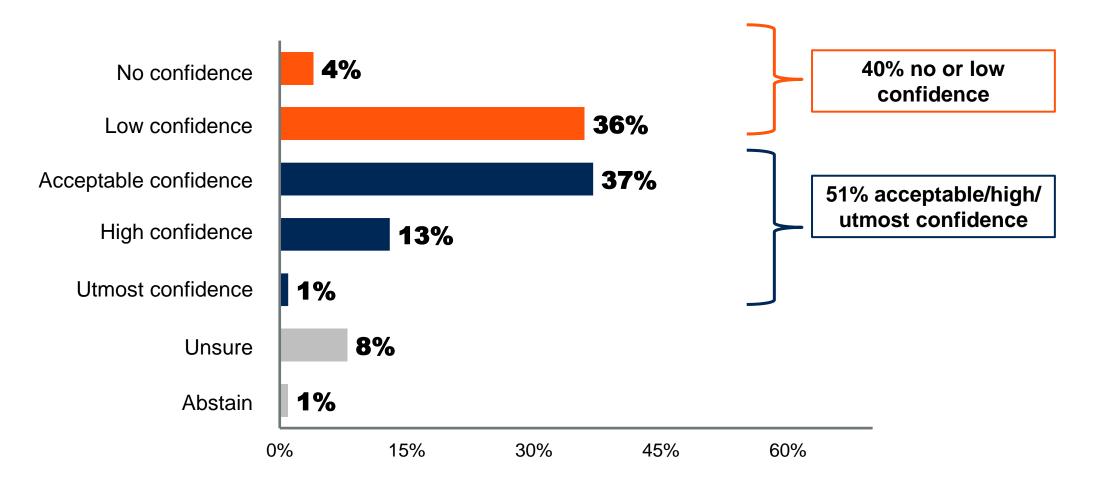
Q: What is the level of confidence you have in your organization's approach to **identifying/assessing risk** for AI initiatives (use cases)? Percentage of respondents





Levels of Confidence: Measuring ROI/Benefits for AI in the Enterprise

Q: What is the level of confidence you have in your organization's approach to **measuring ROI or benefits** for AI initiatives (use cases)? Percentage of respondents





Polling Question: Do you have high confidence in your ability to calculate costs for Al in the **Enterprise?**

- Responses
 - Low
 - Medium
 - High



Polling Question: Do you have high confidence in your ability to estimate, measure and realize value for AI in the Enterprise?

- Responses
 - Low
 - Medium
 - High



Overall Framework for Assessing the Value and Cost of the Eight Use Cases in This Research



Strategically Assess Value

Defend

Competitive parity

Marginal gains/ incrementalism/ microinnovations

Return on employee

Extend

Competitive advantage

Growth in either market size, reach, revenue or profitability

Return on investment

Upend

Market maker

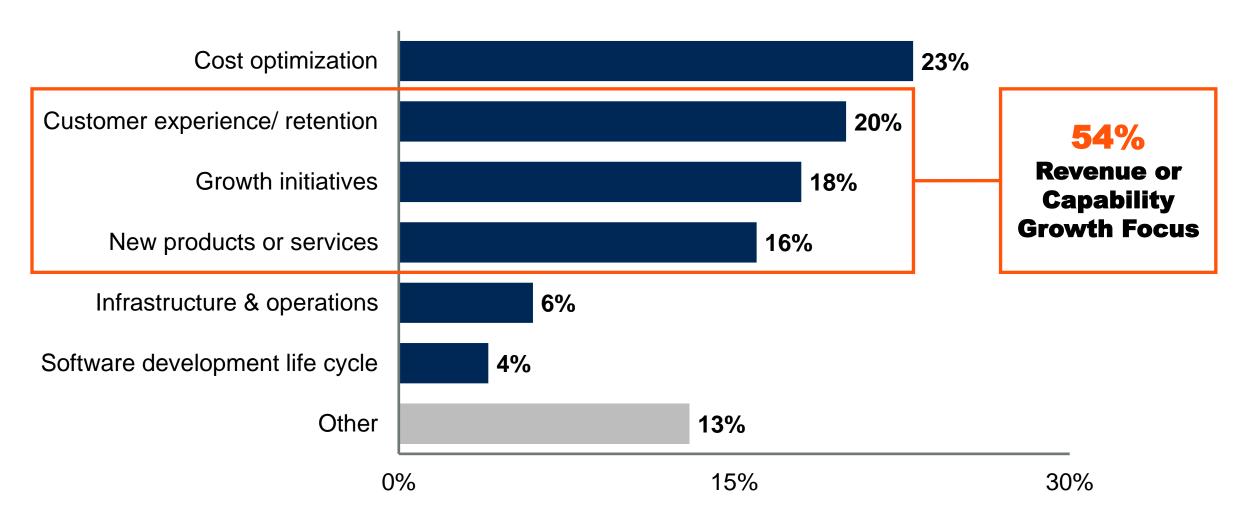
Game changing/ change the competition

Creation of new markets and products

> **Return on** the future



Primary Business Focus of GenAl Initiatives

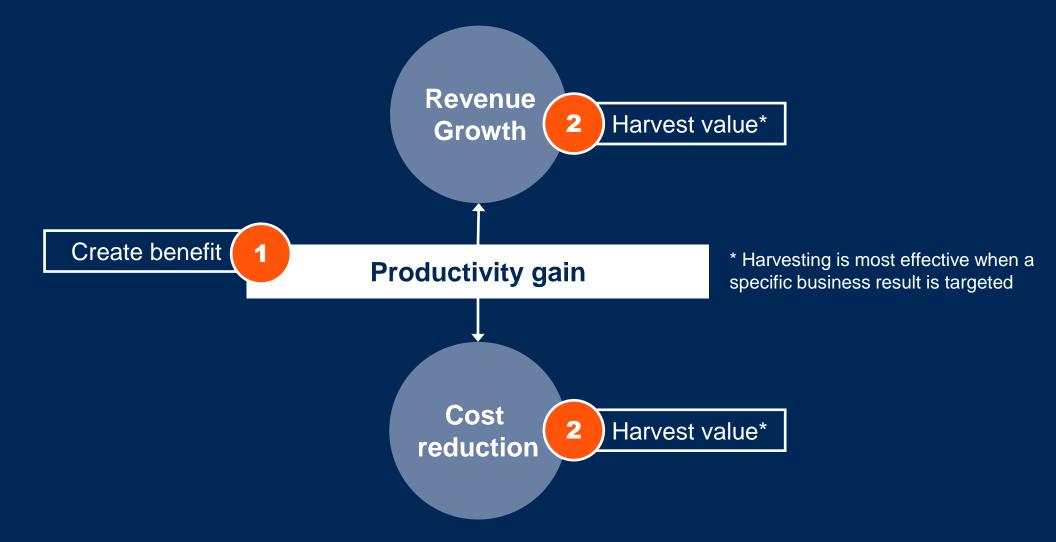




Q: What would you consider your organization's PRIMARY focus for your initiatives? Source: Generative AI Realities: Measuring and Quantifying Business Results Webinar Polling January 2024



'Productivity' Isn't the Same as Value





Productivity Leakage





This Research Calculates Value and **Cost for Eight Common Use Cases**



Defend task-specific improvement, competitive parity

Return on employee

- - Marketing copy generation

• Business assistants/copilots

Coding assistants



Extend an existing process, for differentiation

Return on investment

- GenAl in customer support applications
- GenAl in sales applications
- GenAl document search and summary



Upend to create new products, markets, and core processes

Return on the future

Industry or domain GenAlenabled applications:

- Fine-tuned
- Custom foundation models

GenAl Total Cost, Complexity and Risk





Low

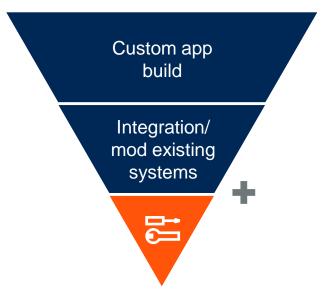
Low

High 4

Value and Competitive Impact

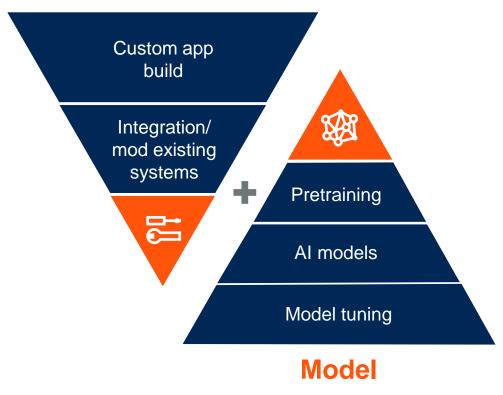


Application

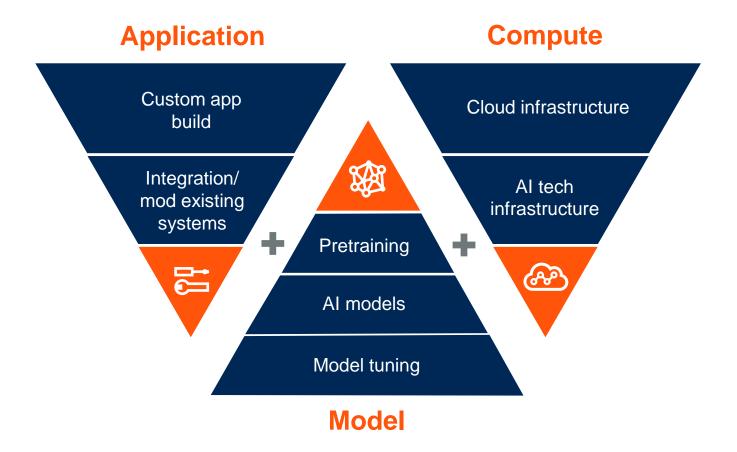




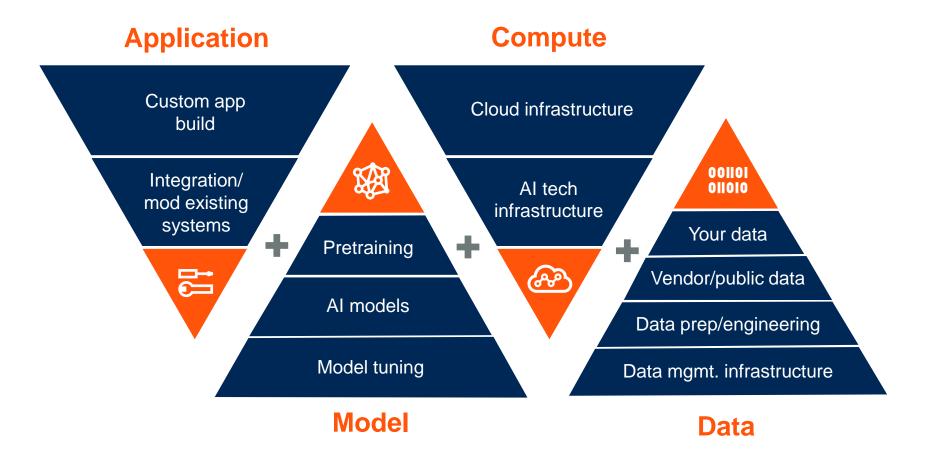
Application



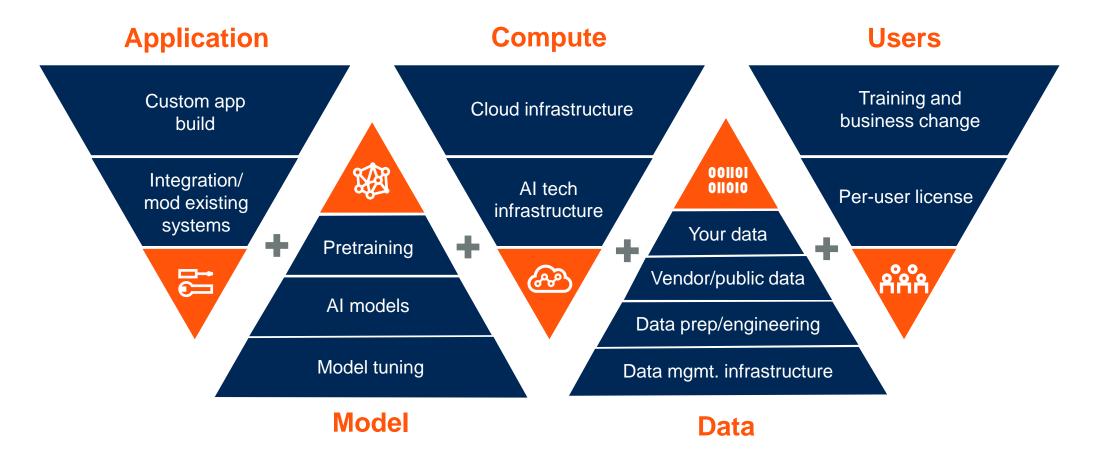




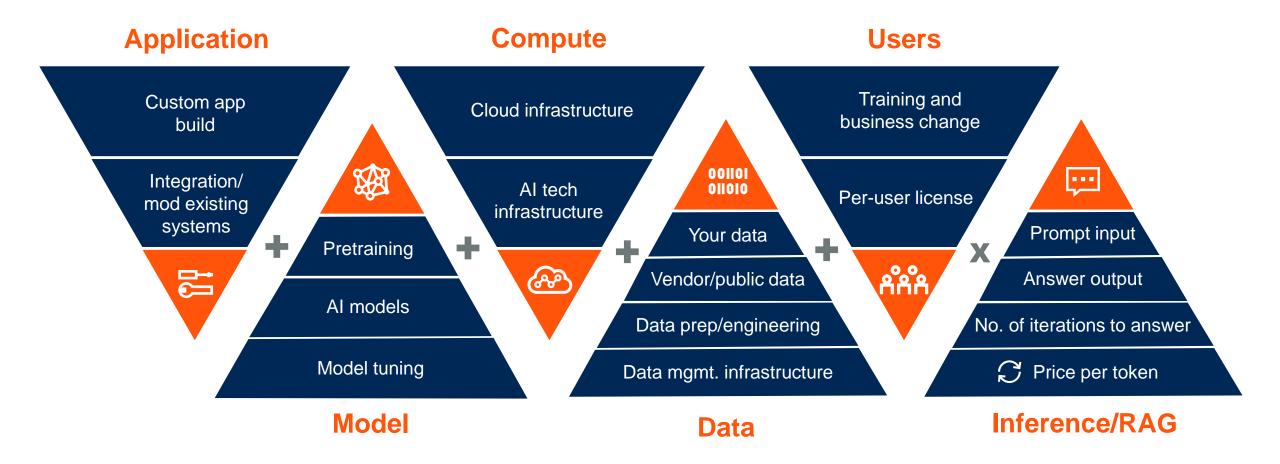






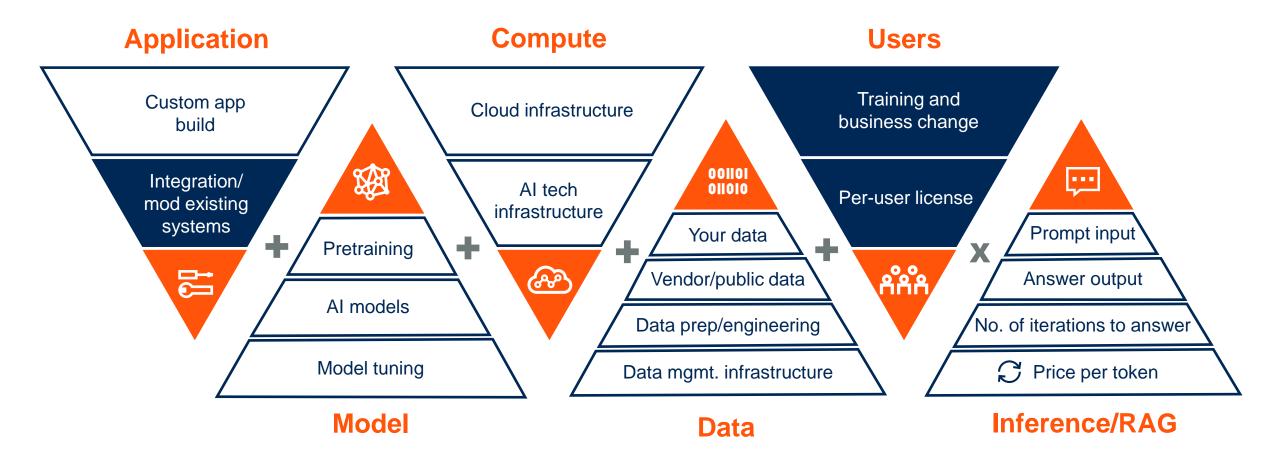








Defend





DEFEND Use Cases: Estimated Total Cost Of Ownership

Total Cost of Ownership Components	Coding Assistants	Business Productivity Assistants	Marketing Content Creation
Initial Pilot and	~\$115K-\$166K includes:	~\$730K-\$800K includes:	~\$450K-\$750K includes:
Rollout, Development, Deployment, Integration, Training	 Developers, platform engineering, security, risk and governance for 3 to 4 months for 200 users User training and business change 	 Office365 specialists doing SharePoint optimization, risk management, security, data engineers, product management, governance and audit User training and business change 	 Developers, data engineers, data scientists, product management, and security, risk and governance working for 3 months for developers; 6 to 12 months for data engineers User training and business change
Recurring Costs	 \$300-\$600 per user per year includes: SaaS application pricing and 10% to 15% of initial deployment costs for 200 users No incremental AI and data management licenses 	 \$200-\$500 per user per year includes: SaaS application pricing and 10% to 15% of initial deployment costs for 1,000 users No incremental AI and data management licenses 	 *\$1.1K-\$3.3K per user per year includes: SaaS application pricing and maintenance cost of 10% to 15% of initial deployment costs for 100 users Al and data management licenses

Source: See the download spreadsheet for details.



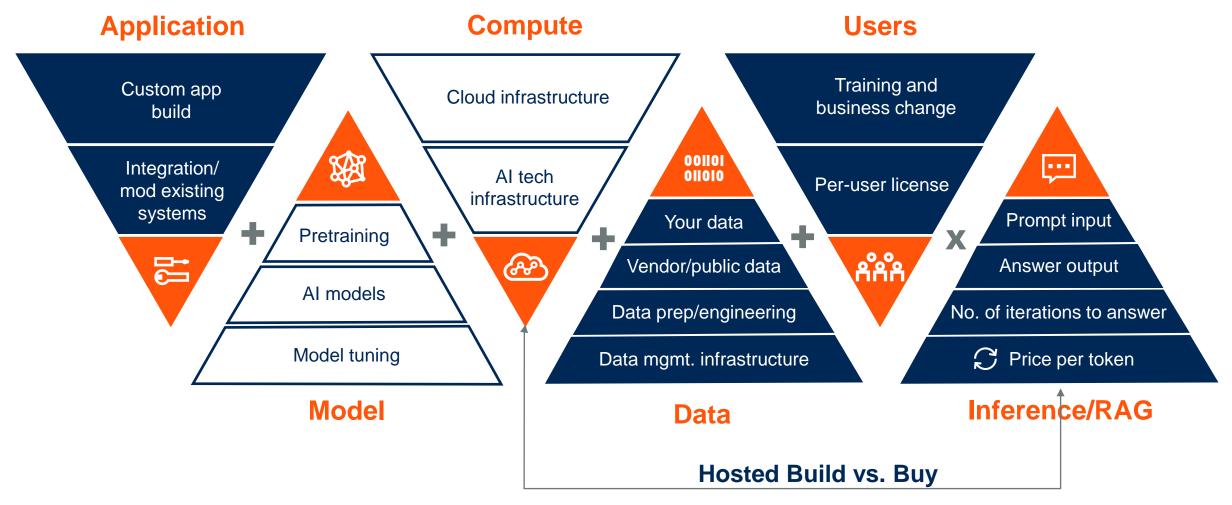
DEFEND Use Cases: Estimated Value

Value Components	Coding Assistants	Business Productivity Assistants	Marketing Content Creation
Value	~\$1.7K-\$3.2K per user per year	~\$1.6K-\$4.1K per user per year	~\$6.5K-12.3K per user per year
	 7%-15% productivity improvement on 25% of time spent writing code with a productivity leak of 20-30%. 	 5-15% productivity improvement, assuming 30% of time spent on creating business documents, with a productivity leak of 20-30% 	• 7-15% productivity improvement, assuming 90% of time spent on marketing content creation with a productivity leak of 20-30%.
Sample KPIs Impacted	 Time to onboard new hires Stories or story points per sprint Stories completed Deployment frequency Number of changes having a customer impact Number of commits Change lead time Number of changes per week Pull request iteration time Lines of code per developer Developer retention Time to upskill junior developers Ratio of junior staff to senior staff Average salary per developer Time to proficiency/onboard 	 Time saved to create business documents Employee retention Percentage of time spent with customers Quicker time to proficiency Time spent on email per day Time to create content Time spent analyzing data Return-on-employee metrics 	 Content output per marketer — do more with the same resources Quality of output — time for rework Average skill level per marketer —junior marketers perform at expert levels Average time to create videos, blogs and marketing copy Savings in agency fees with insourcing Reduction in overhead expense for translation/localization services, on-site photo shoots, studio time, etc. Asset value gain through reuse Reduction in compliance time and resources Return on ad spend (ROAS) Paid media budget savings through improved ad efficiency

Source: See the download spreadsheet for details.



Extend — GenAl Embedded in an Enterprise Application or a Custom Application Embedding an LLM





EXTEND Use Cases: Estimated Total Cost Of Ownership

Total Cost of Ownership Components	GenAl-Assisted Customer Support — Buy	GenAl-Assisted Customer Support — Build
Initial Pilot and Rollout, Development, Deployment, Integration, Training	 \$760K-\$1.3M includes: Developers, data engineers, data scientists, product management, and security, risk and governance working for 6 to 12 months User training and business change for 500 agents 	 \$1.5M-\$3.3M includes: Developers, data engineers, data scientists, product management, and security, risk and governance working for 6 to 12 months User training and business change for 500 reps
Recurring Costs	 \$1.9K-\$10K per user per year includes: SaaS plus API consumption licenses and varies call volume at the lower and upper end Application and model maintenance of 15% to 20% of initial deployment costs for 500 agents AI and data management licenses 	 *\$1.5K-\$5.3K per user per year includes: API token license fees for low- and high-end call volumes and tokens per call Application and model maintenance of 30% to 60% of initial deployment costs for 500 agents AI and data management licenses

For these assumptions, build becomes more favorable after 500 users.

Source: See the download spreadsheet for details.



EXTEND Use Cases: Estimated Total Cost of Ownership

Total Cost of Ownership Components	Personalized Sales Content Creation	Document Search and Summarization
Initial Pilot and Rollout, Development, Deployment, Integration, Training	 \$990K-\$1.8M includes: Developers, data engineers, data scientists, product management, and security, risk and governance working for 6 to 12 months User training and business change for 500 reps 	 \$1.1M-\$2.4M includes: Developers, data engineers, data scientists, product management, and security, risk and governance working for 6 to 12 months User training and business change for 1,000 users
Recurring Costs	 \$1.3K-\$8.4K per user per year includes: SaaS plus API consumption licenses on upper end Application and model maintenance of 15% to 20% of initial deployment costs for 500 reps AI and data management licenses 	 *\$600-\$1.4K per user per year includes: API consumption fees from commercial model 3 to 1 GenAl iterations per query ratio Three sub API context prompt calls per GenAl iteration Application and model maintenance of 20% to 25% of initial deployment costs for 1000 users Al and data management licenses

Source: See the download spreadsheet for details.



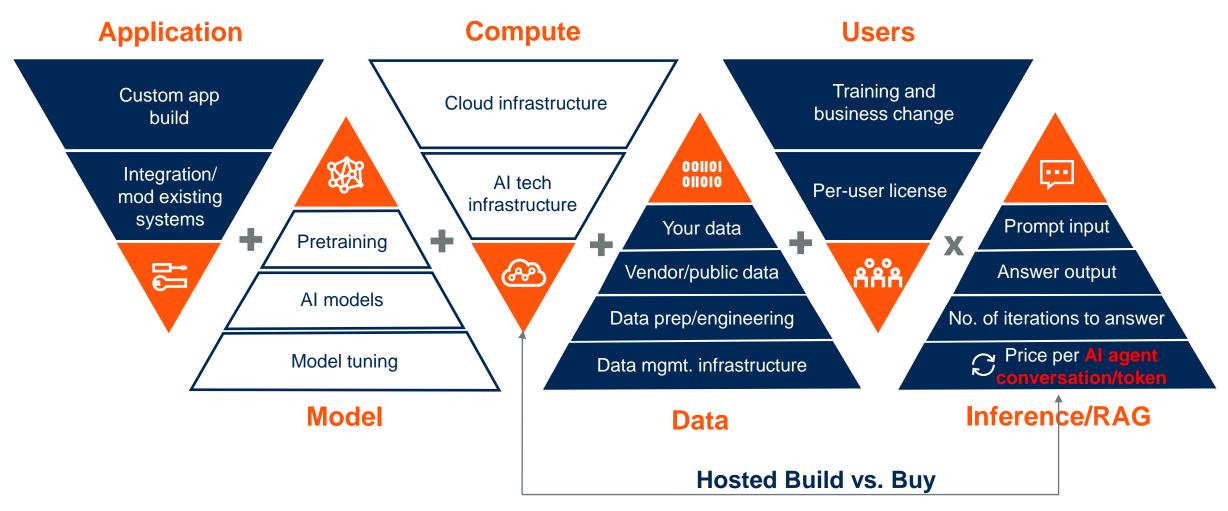
EXTEND Use Cases: Estimated Value

Value Components	GenAl-Assisted Customer Support	Personalized Sales Content Creation	Document Search and Summarization
Value	 *\$6K-\$13K per agent per year 14%-35% productivity improvement, with a productivity leak of 20-30% 	 *\$1.7K-\$4.3K per year per rep incremental contribution to EBITDA 14% to 25% productivity improvement on 6 hours per week spent on creating sales content, with a productivity leak of 20% to 30%, with 50% of saved time used for selling and EBITDA range of 20.3% to 33% 	 *\$1.5K-\$2.1K per knowledge worker per year 7% to 10% productivity improvement, with a productivity leak of 20% to 30%, assuming 30% of time is spent on searching and summarizing information
Sample KPIs Impacted	 •Time to resolution •Response time •Net Promoter Score (NPS) •Call satisfaction scores •Agent retention •Productivity of junior staff •Time to reach expert skill level/onboard new hires •Time to productivity •Number of escalations •Manager's time to train new staff •Incremental revenue from upsell/crosssell •Level of detail of call center analytics (topic ID on all calls rather than a sample) 	•Revenue per sales rep — higher sales productivity leads to more deals per rep in the funnel •Average deal size •Percentage of cross-sell/upsell revenue •Size of pipeline •Close rate •Customer retention •Cost to acquire •Customer satisfaction	Percentage of knowledge worker time spent on the highest-value tasks of their role (respond to customers, build products, sell, and manage suppliers, contracts, claims and legal documents) Time to answer client questions Content quality/error rates

^a Gartner Seller Time Spend Assessment.

Source: See the download spreadsheet for detailed costs.

Extend — Al Agent





EXTEND Use Cases: Estimated Total Cost of Ownership

Total Cost of Ownership Components	Al Agent Plus Human GenAl-Assisted Support — Buy	Al Agent Plus Human GenAl-Assisted Support — Bulld
Initial Pilot and Rollout, Development, Deployment, Integration, Training	 \$1.5M-\$2.7M includes: Developers, data engineers (double the GenAlassisted customer support "buy" use case), data scientists, product management, and security, risk and governance working for 6 to 12 months User training and business change for 500 reps 	 \$2.5M-\$4.4M includes: Developers, data engineers, data scientists, product management, and security, risk and governance working for 6 to 12 months User training and business change for 500 reps
Recurring Costs	 *2.3K-\$15.2K per human and AI agent equivalent per year of 500 includes: 10% low to 30% high-end conversations shifted to AI agents 30 low to 50 high-end calls per agent per day; 20 messages per conversation 50% of messages GenAI 	 *2K-\$6.6K per human and Al agent equivalent per year of 500 includes: 10% low to 30% high-end conversations shifted to Al agents 30 low to 50 high-end calls per agent per day 20 messages per conversation 50% of messages GenAl
	 \$0.30 low price per conversation to \$2.00 high-end price per conversation Application maintenance of 15%- 20% of initial development costs 	 \$0.02 low price per conversation to \$0.065 highend price per conversation. Application maintenance of 30%- 60% of initial development costs

For these assumptions, build becomes more favorable after 500 users.

Source: See the download spreadsheet for details.



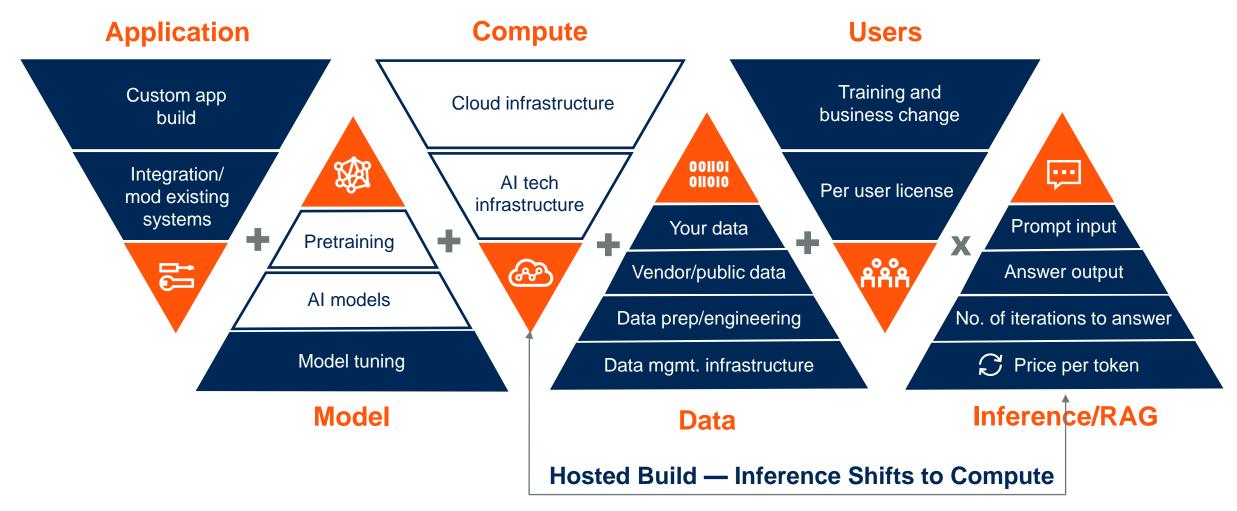
EXTEND Use Cases: Estimated Value

Value Components	Al Agent GenAl-Assisted Support- Buy	Al Agent GenAl-Assisted Support — Build	
Value	 *\$11.4K-\$27.3K per agent per year 10% low to 30% high replacement of human agent with AI agents, plus a ~14% to 35% productivity improvement for remaining human agents, with a productivity leak of 20% to 30. 		
Sample KPIs Impacted	•Time to resolution •Response time •Net Promoter Score (NPS) •Call satisfaction scores •Agent retention •Productivity of junior staff •Time to expert skill level/on board new hires •Time to productivity •Number of escalations •Manager's time to train new staff •Incremental revenue upsell/cross-sell •Level of detail of call center analytics (topic ID on all calls rather than a sample)		

Source: See the download spreadsheet for details.

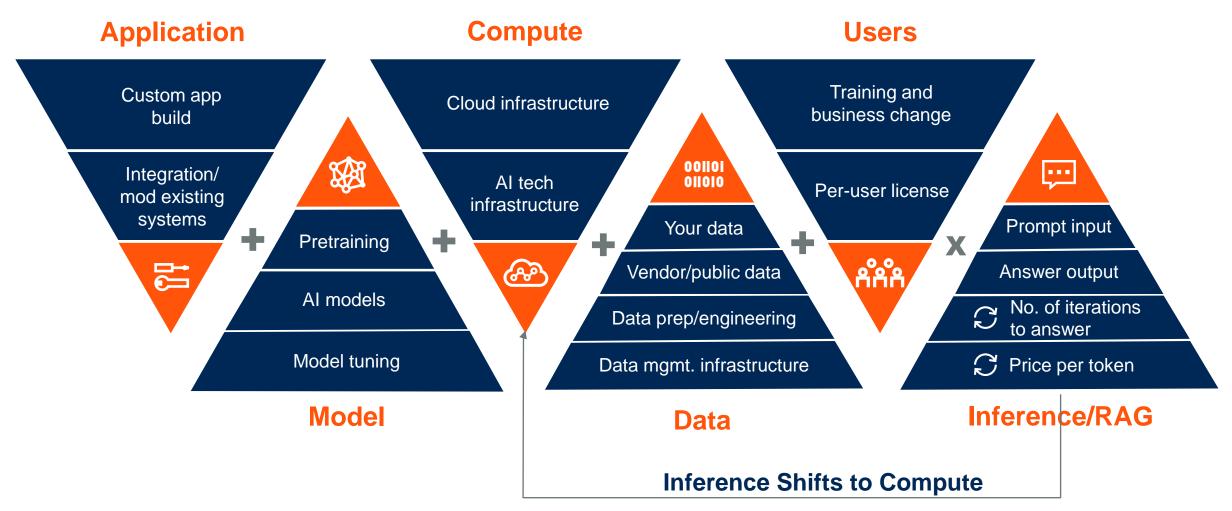


Extend — Custom Application Embedding Fine-Tuned Model





Upend — Fine-Tuned or Custom Model





UPEND Use Cases: Estimated Total Cost of Ownership

Total Cost of Ownership Components	New domain application: fine-tuned SLM (e.g., for insurance underwriting)	New domain application: fine-tuned LLM (e.g., for legal or medical assistants)	New domain application: custom LLM (e.g., for drug discovery)
Initial Pilot and Rollout,	~\$3.5M-\$5.1M includes:	~\$5.1M-\$7.5M includes:	~\$12M-\$16.6M includes:
Development, Deployment, Integration, Training	 Initial fine-tuning of pretrained open-source 8B parameters High-end developers; data engineers; data scientists; security, risk and governance; product management working for 9 to 12 months User training for 1000 users 	 Costs include initial fine-tuning of pretrained open-source 70B parameters High-end developers; data engineers; data scientists; security, risk and governance; product management working for 9 to 12 months User training for 1000 users 	 Model pre-training of a 13B parameter model High-cost developers, data engineers, data scientists, security, risk and governance, product management working for 12 to 15 month User training for 1000 users
Recurring Costs	 \$1.4K-\$2.5K per user per year includes: Ongoing LLM inference costs per GenAl query GPUs for embeddings, application and model maintenance 25% to 30% of initial deployment costs for 1,000 users With 10% (low)- 20% (high) user concurrency and 5,000 (low) to 10,000 (high) tokens per query, plus Al and data management licenses 	 \$2.6K to \$5.2K per user per year includes: Ongoing LLM inference costs per GenAl query GPUs for embeddings, application and model maintenance 25% to 30% of initial deployment costs for 1,000 users With 10% (low)- 20% (high) user concurrency and 5,000 (low) to 10,000 (high) tokens per query, plus AI and data management licenses 	 \$5.6K-\$11.1K per user per year includes: Ongoing LLM inference costs per GenAl query GPUs for embeddings, application and model maintenance of 30%to 40% of initial deployment costs for 1,000 users With 10% (low)- 20% (high) user concurrency and 5,000 (low) to 10,000 (high) tokens per query, plus Al and data management licenses

Source: See the download spreadsheet for details.



UPEND Use Cases: Estimated Value

	New domain application: fine-tuned SLM (e.g., for insurance underwriting)	New domain application: fine-tuned LLM (e.g., for legal or medical assistants)	New domain application: custom LLM (e.g., for drug discovery)
Value	Use-case-specific	Use-case-specific	Use-case-specific
Sample KPIs Impacted	 Market share Percentage of revenue from GenAl products Percentage of underwriting losses Size of new market created Reduced time to assess risk from complex mix of contracts and insurance documents Customer retention Customer lifetime value Brand value improvement from being perceived as an innovator Percentage of services that are now free to drive business in others 	 Market share Percentage of revenue from GenAI products For legal, productivity improvements to contract analysis and negotiation, due diligence, discovery, dispute resolution, litigation support, and audit support Percent time and quality improvement in document generation such as contracts, briefs and pleadings Higher-quality diagnosis and treatment; better healthcare outcomes Reduced error rates in medicine and law 	 Size of new markets created from new drug formulations Revenue from new drug formulations Reduce time for clinical testing of new drugs Time to market for new drugs Time to predict risks and opportunities in cases

Source: See the download spreadsheet for details.



As GenAl models become increasingly commoditized, realizing competitive advantage and full benefits from GenAl will be driven by how effectively an organization uses unique data, redesigns work/processes, and manages change and risk.

What Is Your Al Ambition?

Defend

...task specific improvement to maintain competitive parity

Microsoft Copilot, Google Gemini, Apple Intelligence

> Incrementalism Marginal gains **Microinnovations**

Individual

Extend

... an existing process for differentiation

GenAl in customer service to recommend upsell

Growth in market size, reach, revenue or profitability

Process

Upend

... to create new products, markets and core processes

> **Pharmaceutical** drug discovery

Change the business model and competition

Business Model

Gartner

Strategically Plan And Actively Harvest Value

Defend

Competitive parity

Marginal gains/ incrementalism/ microinnovations

Manage like a cost center

Extend

Competitive advantage

Growth in either market size, reach, revenue or profitability

Manage like a profit center

Upend

Market maker

Game changing/ change the competition

Creation of new markets and products

Manage like a venture capital port



Imperative: Al demands that org. defines, develops and curates a portfolio.

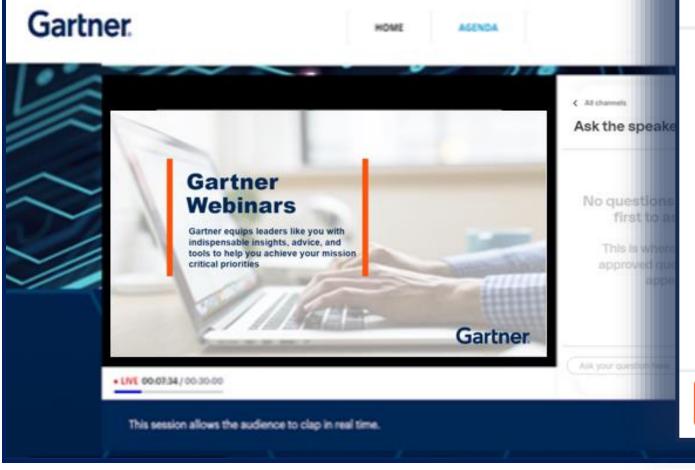


Recommendations

- Measure and manage total cost and value from POC to enterprise scale based on whether use cases are 'defend', 'extend' or 'upend'
- Don't underestimate the need to invest in AI ready data and governance
- Don't underestimate the need to invest in and manage business change



Ask the speaker



All channels Ask the speaker No questions yet. Be the first to ask one! This is where all of the approved questions will appear

Ask your question here





Gartner CIO Leadership Forum

6 – 7 May 2025 I Sydney, Australia

9 – 10 July 2025 | Tokyo, Japan

Gartner CIO Leadership Forum is an unrivaled experience — a place where you'll access the most timely advice, actionable tools and CIO peer insights to adapt faster and stay ahead of change.

Learn more: gartner.com/conf/cio

At this year's forum, you'll learn:



Short and long-term strategies to **lead you** and your team to success in dynamic times.



Reshape organizational structures and attitudes for greater agility.



Engage peers to co-create ways of working that remove digital friction to speed business outcomes.



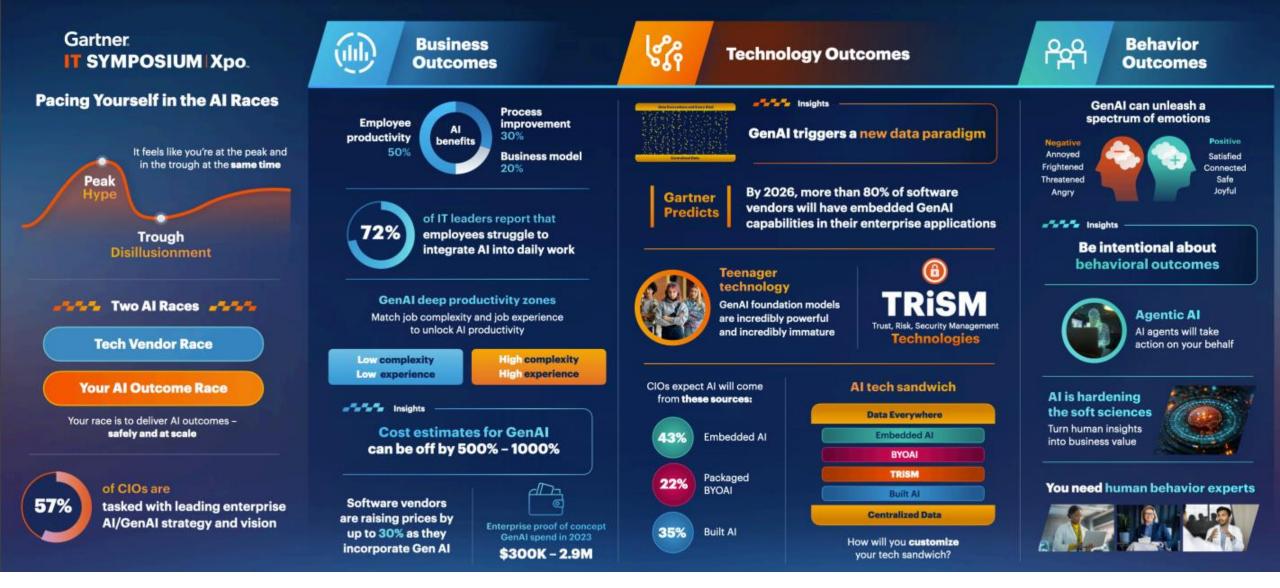
Get Al Ready — What IT Leaders Need to Know and Do

Ready your enterprise to capture Al opportunities and bolster your cybersecurity, data and Al policies and principles.



Download Resource





Get the guidance you need to set the right AI pace and safely achieve your business, technology, and behavioral goals at scale.

Gartner for IT on social media

Want to stay in-the-know? Connect with us on LinkedIn and Twitter to receive the latest Gartner IT insights and updates across research, events and more. It's all curated specifically for IT leaders and decision-makers.

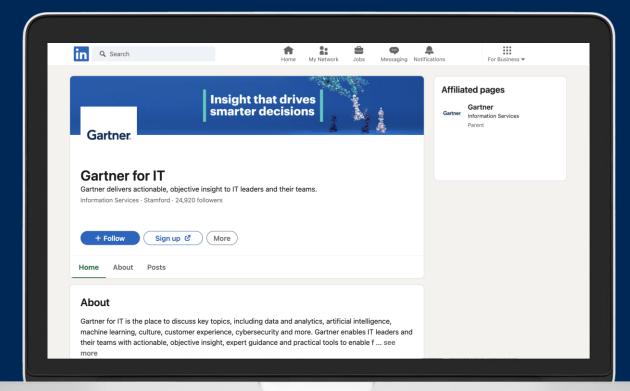
Follow us on





Looking for insights delivered to your inbox?

Subscribe to our



Gartner

Become a Client

Clients receive 24/7 access to proven management and technology research, expert advice, benchmarks, diagnostics and more.

Fill out the form to connect with a representative and learn more.

Learn More

Or give us a call: +441784614280 | +1 855 637 0291

8 a.m. – 7 p.m. ET 8 a.m. – 5 p.m. GMT Monday through Friday



Get more Gartner insights



Download the research slides



View upcoming and on-demand Gartner webinars at gartner.com/webinars



Rate this session



Rate this session

