WHITEPAPER

Build Generative AI Literacy to Embrace the Future of Knowledge Work

TABLE OF CONTENTS

INTRODUCTION

The nature of knowledge work is undergoing a fundamental shift. Today's professionals must navigate a relentless flood of unstructured and fragmented information that is difficult to access effectively. Technology leaders face a dual mandate:

- **Keep pace with accelerating complexity.**
- **Strategically harness new capabilities that turn information into an actionable advantage.**

Deployed with purpose, generative AI can transform how organizations capture, retrieve, and apply knowledge. It enables a shift from static information systems to dynamic engines of insight. But realizing this potential requires more than adopting tools that promise immediate outcomes. True advantage will belong to organizations that build **Generative AI Literacy** across the workforce to ensure domain experts understand how these systems function, where they add value, and where human oversight is still essential.

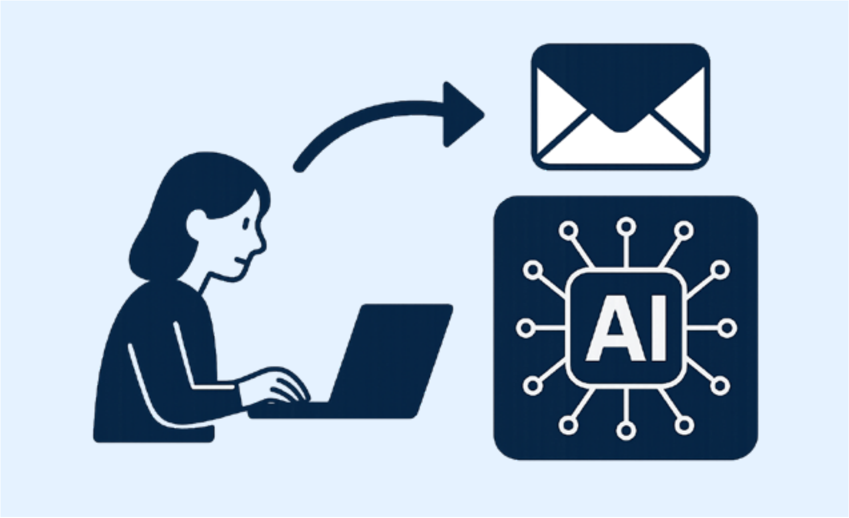
This paper offers a practical foundation for building that literacy. It begins by making the case for Generative AI Literacy in modern organizations and introduces three core pillars: Awareness, Confidence, and Readiness. It then explores the first of these pillars – Awareness – in more depth by laying out essential concepts knowledge workers should understand to guide their interactions with AI systems. The paper concludes with a discussion on building organizational confidence in generative AI through practical experience and operationalizing generative AI in the flow of knowledge work.

These concepts are not all you should know about how generative AI systems work, but they should offer a solid starting point. Organizations that invest early in building solid understanding of generative AI across their workforce to pair with their earned domain knowledge should be best positioned to lead the future of knowledge work.

THE CASE FOR GENERATIVE AI LITERACY IN KNOWLEDGE WORK

Discussions around artificial intelligence often swing between two extremes: enthusiastic optimism about generative AI's potential, and cautious skepticism about its risks and limitations. Somewhere between these extremes lies the practical reality that generative AI systems are already reshaping knowledge-intensive industries - from financial services and healthcare to government and other highly regulated industries.

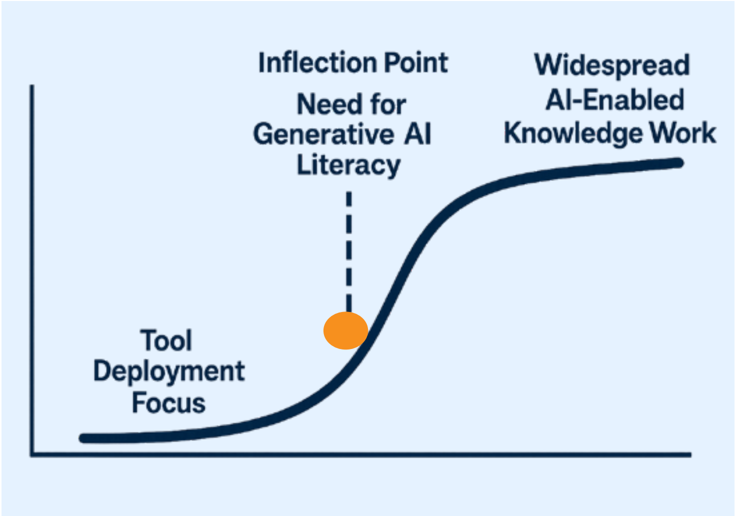
PROOF POINT: GENERATIVE AI IN ACTION



Real-world examples underscore both the scale of this shift and the growing need for deeper understanding. For example, Allstate has implemented generative AI models (specifically OpenAI's GPT models) enhanced with company-specific terminology to improve customer communications. These AI-drafted emails have proven more consistently empathetic and less jargony than human-written messages, leading to significantly improved claimant interactions. Today, the majority of Allstate's 50,000 daily communications from its 23,000 representatives to claimants are written by AI systems and verified by humans for accuracy (Mattioli, 2024).

This example underscores a critical requirement: knowledge workers must provide domain-specific context to these systems (in this case evolving company-specific policies and terminology) and remain engaged in applying their knowledge by assessing and validating outputs.

INFLECTION POINT



The earlier example also marks a broader inflection point - success with generative AI does not hinge solely on technical deployment. It increasingly depends on human teams developing Generative AI Literacy - the ability to guide, evaluate, and work effectively alongside these systems. Domain knowledge is essential, but so is an understanding of how generative AI interprets and applies that knowledge to produce quality outcomes.

Despite significant investment in tooling, knowledge workers continue to spend a substantial part of their day searching for information across fragmented systems. According to the International Data Corporation (IDC), knowledge workers spend roughly 2.5 hours per day (or 30% of their time) just finding information (ProProfs, 2023). Emails, reports, contracts, presentations, customer interactions. All hold potentially valuable insights yet stay underused because tools alone (even generative AI tools) cannot surface insight reliably without structured understanding and human context.

BRIDGING THE GAP

As generative AI adoption expands, what's needed is not just wider access to tools, but deeper Generative AI Literacy: a working understanding among knowledge workers of how these systems retrieve, process, and generate content. This literacy must complement, not replace, the deep domain knowledge employees already bring to their roles.

Generative AI offers a powerful path to bridge the gap between vast unstructured information and practical knowledge. Still, there is a gap between public enthusiasm and internal, practical understanding of how these systems work. Often knowledge workers continue to view AI as a "magical" black box, rather than a logical tool shaped by data, training methods, and design decisions. Yet when deep domain knowledge is paired with evolving Generative AI Literacy, the results can be transformative.

GENERATIVE AI LITERACY PILLARS

Building Generative AI Literacy requires focus across three essential pillars: Awareness, Confidence, and Readiness. These pillars directly address the most persistent barriers organizations face when integrating generative AI into knowledge work: lack of foundational understanding, mistrust or misuse of AI systems, and challenges in operationalizing AI within real workflows.

- **Awareness:** Equipping knowledge workers with a foundational understanding of how generative AI systems are trained, how they retrieve and process input, and why context and prompting strategies matter. Without this baseline, knowledge workers risk misinterpreting generative AI outputs or using systems ineffectively.
- **Confidence:** Building trust through practical, firsthand experience and transparent conversations about strengths, limitations, and responsible use. Confidence empowers employees to ask better questions, challenge results appropriately, and set clearer expectations around generative AI-assisted work.
- **Readiness:** Offering structured opportunities for experimentation, guided learning, and cross-functional collaboration to integrate AI into daily workflows. Readiness ensures that AI tools are not isolated pilots but are embedded within core knowledge processes.

These pillars form a practical framework for bridging the gap between technological capability and human knowledge. Awareness ensures that knowledge workers understand how generative AI systems function and it influences how they interact. Confidence ensures they trust the systems enough to use them critically and effectively. Readiness ensures they are prepared to integrate generative AI into real-world tasks and decision-making.

Building generative AI literacy is no longer an optional endeavor for technology leaders, it is a strategic imperative. Organizations that invest in developing practical AI understanding across their workforce will not only improve tool adoption but will also unlock the potential of generative AI - positioning themselves to lead the next era of knowledge work with agility, confidence, and precision.

AWARENESS: UNDERSTAND GENERATIVE AI IN CONTEXT

Building Awareness, the first pillar of Generative AI Literacy, begins with clarifying concepts that are often misunderstood even among experienced knowledge workers and leaders. LinkedIn's 2024 Work Change Report notes a 177% increase in members adding AI literacy skills, which indicates people's growing recognition of AI's importance (LinkedIn, 2024). But understanding these core concepts is not simply an individual or academic exercise; it can shape the way organizations engage with generative AI tools in real-world environments.

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

One of the most common sources of confusion is the interchangeable use of the terms Artificial Intelligence and Machine Learning. Understanding their relationship is a good place to start.

Artificial Intelligence (AI) refers broadly to systems designed to simulate human intelligence, such as pattern recognition, decision-making, and problem-solving.

Machine Learning (ML) is a subset of AI focused specifically on enabling systems to learn directly from data based on past or representative examples, rather than relying solely on explicitly programmed rules.

Think of it this way; AI is the goal and Machine Learning is one of the key methods used to achieve it. AI is the broader field concerned with making machines "smart," and ML is an important approach we use to train those machines to behave intelligently.

Here's a simple illustration. A traditional AI system in customer service might follow a precise decision tree:

- **If the customer says, "reset password," then offer reset instructions.**
- **If the customer says, "close account," then escalate to a representative.**

This rule-based system is technically AI because it is simulating the way a human may respond to a command, but it is rigid and only responds to predefined inputs. Alternatively, a machine learning-based system could be trained on thousands of real customer conversations. It won't just match exact phrases; it will recognize patterns from these examples and can "learn" to recognize intent without explicit rules.

This is the distinction that matters; not all AI involves learning from data, but all Machine Learning does. And modern generative AI models are advanced machine learning models. Recognizing this provides an important framing how generative AI systems evolve, adapt, and sometimes fail depending on how knowledge workers interact with them, and the quality of context (data) they absorb.

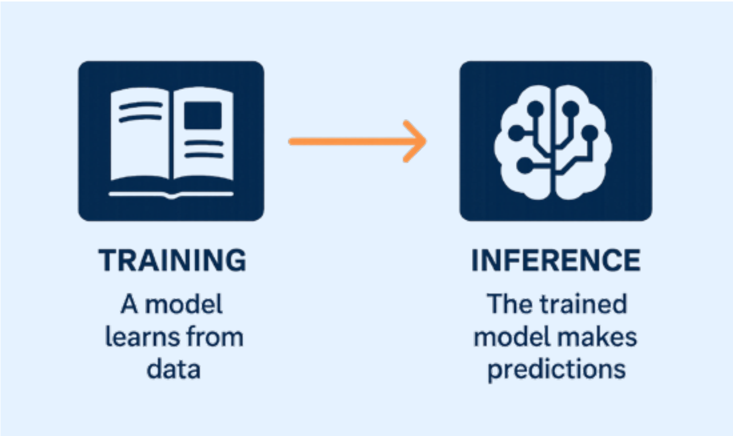
CATEGORIES OF ARTIFICIAL INTELLIGENCE

Diving deeper, Machine Learning models themselves can be categorized into two major approaches: discriminative and generative models. For the purposes of this discussion, we'll refer to the to them with the broader "AI" label which you'll hear more commonly. But know that in this part of the discussion and going forward in this section, we are talking about advanced machine learning models under the broader Artificial Intelligence umbrella.

- **Discriminative AI models** are designed to recognize boundaries between categories or outcomes. For example, a discriminative AI system might be trained to distinguish between spam and legitimate emails. They classify inputs based on differences learned from earlier examples.
- **Generative AI models** move beyond categorization. They learn the underlying distribution of data so well that they can interpret intent or create entirely new examples that resemble the original data. For instance, a generative AI system trained on thousands of email threads about product inquiries could generate entirely new, draft responses to similar future product questions. These responses are not copied from the training data, but composed based on learned tone, structure, and intent.

This distinction is important for knowledge workers to understand as they are selecting tools and platforms. Where discriminative AI systems can label documents or classify emails, generative AI systems can summarize reports, propose insights, and even draft original content based on organizational knowledge. And taking it one step further, because they can understand intent more fluidly, often generative AI systems are used to orchestrate interaction with discriminative models without the need for firsthand direction from a user.

TRAINING AND INFERENCE



Another concept for building Awareness is training versus inference:

Training is the resource-intensive process where a model learns from existing datasets and then models the underlying patterns it discovers. Think of it like instructing a student. The AI model is given books, articles, images, conversations, and just like a student, it learns patterns, relationships, and structures from that data.

Inference occurs when a user interacts with a trained model. New input is processed against the patterns learned in training to generate predictions, retrieve content, or create new outputs. Inference is like that student now taking an exam or drafting an essay based on what they've learned.

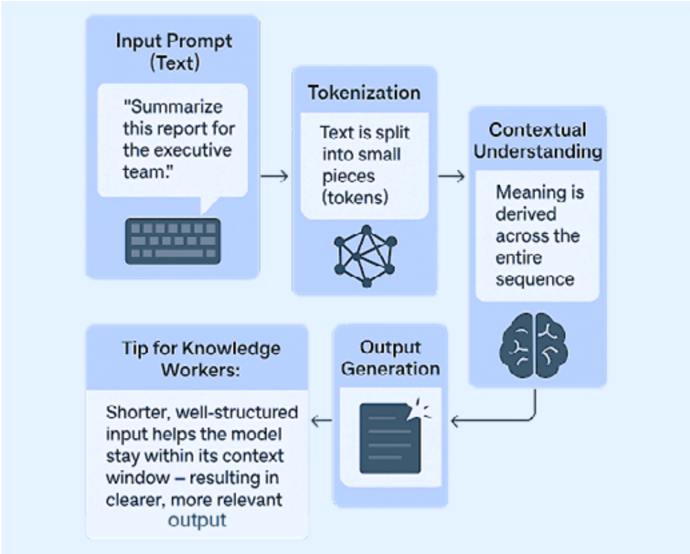
For a more concrete example of inference, asking a generative AI system like ChatGPT a question triggers inference. The rapid, contextually informed response is not the result of new learning happening in real time. It is an application of knowledge accumulated during prior training. Unlike training, inference needs to be fast and efficient because it happens every time we interact with AI.

So why is it important to understand training vs. inference? Because as knowledge workers, we won't be routinely training large AI models – we'll be using them in the inference phase. Understanding inference will help knowledge workers:

- **Choose the right AI models (do you need maximum speed or maximum accuracy?)**
- **Shape better prompts to get higher-quality results.**
- **Organize your data effectively so AI can retrieve and generate meaningful insights.**

Industry is investing heavily in both training and inference, but as generative AI becomes part of everyday workflows, the demand for faster, more efficient inference is skyrocketing. Every time we use AI in our daily work - whether it's summarizing documents, drafting reports, or generating more specific insight - you're relying on inference. And just as we've seen massive investments in training AI models, industry is now focusing on making inference smarter, cheaper, and more accessible so that AI can become a seamless part of how we work.

THE ROLE OF TRANSFORMER-BASED MODELS



As knowledge workers begin to engage more confidently with generative AI, it's helpful to understand the architectural breakthrough that underpins nearly all modern generative AI systems: the transformer. Introduced in a 2017 research paper titled "Attention Is All You Need," the transformer architecture has become the foundation for leading large language models. It's not necessary to understand every detail of how these models work, but developing a working intuition about how they process language can improve how we interact with them.

What makes Transformers Different?

Earlier AI systems struggled to capture long-range meaning in language. They often treated sentences as flat sequences, missing how words relate to each other. Transformers solved this by introducing a mechanism called self-attention, which allows the model to weigh the importance of each word in relation to every other word regardless of position.

For example, in the sentence "The customer who called this morning needs a refund," a transformer can recognize that "customer" is the subject of "needs a refund" even though they're separated by several words. This ability to preserve and use context is what makes transformers so powerful for knowledge work tasks like summarization, Q&A, and content generation.

Tokens, Attention, and Context Windows

To understand how transformers work in practice, consider the following concepts:

- **Tokens:** Transformers don't process whole words, they process tokens. Tokens are pieces of words ("understand" might be split into "under" and "stand"). Every prompt is broken down into tokens before processing.
- **Attention:** Through "self-attention," the model evaluates how heavily each token should influence others. This allows the model to keep meaning across long passages, even if words are far apart.
- **Context windows:** A model can only process a limited number of tokens at once this is its context window. If your prompt exceeds that window, earlier content may be truncated, which affects performance and output relevance.

These mechanics matter because they directly affect how knowledge workers interact with generative AI systems. Well-structured, concise input leads to better output. Overloading the system or assuming it "remembers everything" can degrade performance.

Why It Matters for Knowledge Work

Transformer models are not just text predictors. When used effectively, they can:

- **Summarize lengthy documents.**
- **Generate coherent, context-sensitive drafts.**
- **Translate or rephrase technical content.**
- **Help find inconsistencies or gaps in language.**

But when misused through unclear prompts, overly long inputs, or unrealistic expectations they can generate irrelevant or inaccurate content. That's why awareness of model architecture, even at a conceptual level, is valuable: it shapes how we approach prompts, chunk our content, and evaluate results.

From Theory to Practice

Most knowledge workers will never build or train a transformer model. But they will increasingly depend on them in the inference phase through search assistants, policy copilots, documentation summarizers, and workflow agents embedded in everyday tools. By developing a practical understanding of how these systems process information - via tokens, attention, and context – knowledge workers can unlock higher-quality interactions by considering how these models work. This is Awareness in action: applying foundational knowledge to improve outcomes, not just accepting results at face value.

ORGANIZING KNOWLEDGE FOR RETRIEVAL

Transformer-based models are powerful not just because they generate human-like language, but because they understand relationships between words, phrases, and concepts even when the input is messy or inconsistent. This capability is essential for effective use in knowledge work, where information is rarely cleanly labeled or meticulously organized. To make sense of unstructured data at scale - emails, reports, transcripts, case notes - generative AI systems rely on embeddings and vector databases.

Embeddings

Embeddings are how AI systems mathematically express meaning. Instead of treating language as plain text, transformer models convert words, phrases, or entire documents into high-dimensional numerical vectors, points in a multi-dimensional space. These vectors capture the semantic relationships between pieces of information based on how they appear in training data.

Here's a straightforward way to think about it; just as we might say two ideas are "close in meaning," an embedding captures that relationship literally by placing semantically similar content closer together in vector space. For example:

The phrases *"customer onboarding process"* and *"new client setup"* might look different on the surface, but a well-trained model would position their embeddings near each other - recognizing they refer to the same concept.

This ability allows systems to retrieve related content, even when the exact wording doesn't match.

Vector Databases

Once text is transformed into embeddings, it needs to be stored in a system refined for similarity-based retrieval, not only keyword search. That's where vector databases come in. Unlike traditional databases that match exact terms or values, vector databases are designed to answer questions like: "What are the top 5 documents most similar to this question?" They do this by:

- **Storing embeddings efficiently.**
- **Comparing new inputs to stored vectors.**
- **Returning the most related results based on distance in vector space.**

This is critical for Retrieval-Augmented Generation (RAG), enterprise search, chat assistants, and any AI-enabled tool that needs to ground its output in relevant context.

Why it matters for Knowledge Work

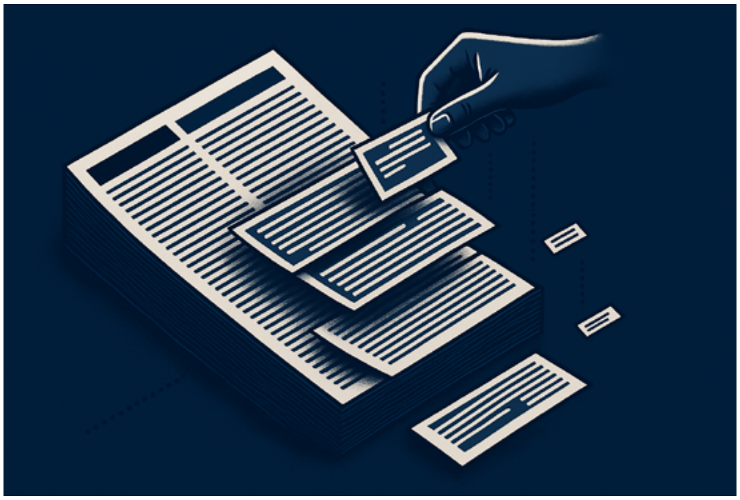
In practice, this means:

- **A policy assistant can surface the right paragraph from a 60-page document even if your question uses different wording than the document itself.**
- **A customer support assistant can retrieve the most relevant product fix even when the issue is described in informal language.**
- **A contract analyzer can group clauses across hundreds of files even if they use different phrasing to express the same legal intent.**

For knowledge workers, understanding embeddings and vector databases is key to improving the structure of source content (writing clearly, using consistent terminology), confirming the quality of references, troubleshooting when AI outputs are poorly grounded or contextually weak. By understanding how information is stored, represented, and retrieved, knowledge workers can partner more effectively with generative AI systems and not just use them passively.

Just as we wouldn't expect a person to recall a concept they were never taught; we shouldn't expect a generative AI system to retrieve insights that weren't represented effectively (embedded). The better we organize and expose our organization's knowledge; the more reliably generative AI can surface this knowledge regardless of how questions are phrased.

MAKING DATA AI-READY WITH CONTENT CHUNKING



As organizations integrate generative AI into more workflows, a common misconception persists; that simply feeding existing content into an AI system will yield meaningful results. In practice, generative AI systems are limited by context windows - the number of tokens (i.e., pieces of words) they can consider at once. If a document exceeds that limit, important context may be dropped or misinterpreted, reducing output accuracy and utility.

A commonly employed approach for addressing this challenge is "content chunking." Content chunking is the process of breaking down information into logically organized, appropriately sized segments. It's not just a formatting exercise, thoughtful chunking yields measurable benefits:

- **Improved retrieval performance: More relevant matches during similarity search.**
- **Faster, more relevant responses: Reduced inference latency.**
- **Greater transparency: Smaller, coherent outputs are easier to confirm and debug.**
- **Better prompting: Clearer chunks lead to better grounding and easier prompt design.**

These advantages are critical not just for improving AI performance, but also for building trust in AI-generated results. Chunking makes it easier to explain, trace, and refine how AI systems arrive at their outputs especially in regulated or high-stakes environments.

Common Chunking Strategies

Several chunking methods are often used in practice, each suited to different content and use cases:

- **Fixed-Length Chunking:** Splits content into standardized sizes (e.g., five hundred tokens per chunk). This is simple to implement and efficient for uniform, repetitive content like news articles or FAQs. However, it risks cutting off ideas midstream—reducing semantic coherence.
- **Context-Aware Chunking:** Uses natural breakpoints like headings, sections, or semantic boundaries to group related ideas. This approach is more complex but better preserves meaning—especially important for contracts, research reports, or policies where precision matters.
- **Hierarchical Chunking:** Combines both approaches by layering summary-level chunks over more detailed sub-chunks. This allows systems to "zoom in" when needed and is effective for large knowledge bases or technical documentation.
- **Multi-Modal Chunking:** Aligns structured and unstructured content (e.g., text with tables or diagrams) into coherent packages. This is critical in fields like healthcare or finance, where relevant insights may span multiple formats.

Each of these strategies balances chunk size, semantic completeness, and retrieval relevance. Choosing the right method depends on content type, user intent, and system architecture.

Why It Matters for Knowledge Work

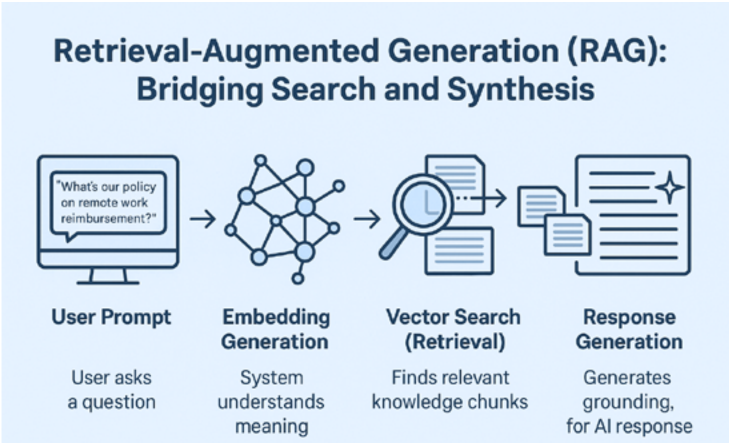
Generative AI systems don't read like humans. They process information in blocks, and once those blocks exceed a certain size, earlier content can be forgotten or ignored. Chunking helps mitigate this by reshaping content into pieces that preserve meaning while staying within the technical boundaries of underlying generative models.

Well-structured chunks improve retrieval by generative AI systems and reduce the risk of hallucinated or incomplete outputs during inference. While generative AI can help parse and organize content, human judgment is still essential. Knowledge workers can play a key role in:

- **Authoring content that is chunk-friendly by using clear sectioning, logical flow, and consistent terminology.**
- **Validating chunks to ensure that meaning, tone, and business context are preserved.**
- **Finding natural breakpoints and semantic boundaries based on domain knowledge.**

This collaboration between machine capability and human structure is what makes chunking effective in real-world applications. Knowledge workers aren't just consuming generative AI as a finished product, they are helping shape how AI interacts with knowledge.

BRIDGING SEARCH AND GENERATION WITH RAG



Generative AI models are powerful content interpreters and generators. But on their own, they don't "know" your organization. Once trained, they can't access new data unless retrained, and their responses are limited by what they've previously seen. In fast-changing environments where content, policies, and context evolve constantly, that's a critical limitation.

Retrieval-Augmented Generation (RAG) addresses this challenge directly. It's a system architecture that combines the semantic search capabilities of vector databases with the language generation power of large language models. The result is hopefully grounded, context-rich outputs that reflect real-time organizational knowledge, not just what the model was trained on.

At a high level, RAG systems work in two stages:

- **Retrieve:** Given a user prompt, the system first converts that input into an embedding, then searches a vector database to retrieve the most semantically relevant chunks of information.
- **Generate:** These retrieved chunks are then passed as context into a generative model, which builds a response grounded in the retrieved content.

This structure allows the generative AI system to reference current, domain-specific knowledge (documents, policies, playbooks, reports) without needing to encode everything into the base model itself.

Connecting the Core Concepts

RAG doesn't stand alone - it brings together the core building blocks explored earlier in this paper. This interconnected workflow allows organizations to use generative AI without retraining models every time a document changes.

- **Embeddings** convert prompts and source content into semantically searchable vectors.
- **Vector databases** store those vectors and return relevant matches based on meaning, not keywords alone.
- **Content chunking** ensures that the material stored in those databases is segmented for retrieval and generation.
- **Transformer-based models** use the retrieved content as context for generating correct, relevant responses.

Reducing Hallucination and Improving Trust

One of the most common concerns with generative AI is hallucination – a situation that occurs when the model generates content that sounds confident but is factually incorrect. RAG reduces this risk by providing source-grounded input at the time of generation. Outputs of RAG often include traceability in the form of references back to the documents that shaped the answer to create transparency for end-users. For high-stakes domains like financial services, healthcare, government, this is essential. It allows generative AI to improve business processes without undermining accuracy or trust.

Why It Matters for Knowledge Work

RAG is particularly well-suited for knowledge-heavy domains where answers need to be correct, current, and grounded in specific organizational content.

Examples include:

- **Internal policy assistants:** surfacing approved guidance based on live documentation.
- **RFP co-authors:** drafting responses that reflect a company's actual project history.
- **HR support tools:** answering employee questions using current benefits and PTO policies.
- **Legal research copilots:** summarizing relevant case law or contract clauses with traceable references.

In each of the examples, what's produced is not just speculative. The output is anchored to real content provided by semantic retrieval. For knowledge workers, RAG isn't just a technical concept. It's the foundation of many generative AI tools they already use (or soon will).

Knowing how RAG systems retrieve and ground their responses can improve how employees:

- **Ask better questions.**
- **Provide stronger source content.**
- **Validate system outputs.**
- **Troubleshoot unexpected results.**

And for leaders, RAG is a practical deployment model. It can aid in unlocking the value of generative AI without requiring investment in ongoing training and fine-tuning a generative AI model.

THE RISE OF AI AGENTS IN KNOWLEDGE WORK

As organizations continue to integrate generative AI into core business processes, the evolution from task-based assistants to goal-driven agents is an important development. Early interactions with generative AI have often been centered on one-off prompts; summarize this document, generate that response. But more and more, AI systems are being designed to act more like collaborators capable of executing multi-step tasks, reasoning across tools, and adapting autonomously to shifting goals.

These “AI agents” extend the capabilities of generative AI far beyond conversation or search. Where a typical assistant responds to a user’s query in a single turn, an AI agent can be assigned a goal, decompose it into steps, access tools or data sources, and conduct a sequence of actions with minimal human input.

What makes it an “Agent”?

Generative AI Assistants	Generative AI Agents
Response-driven behavior	Goal-driven behavior
Single-turn reasoning	Multi-step reasoning
Limited tool use	Broad tool use
Simple, narrow tasks	Complex, autonomous tasks

Three characteristics typically define an AI agent:

Goal-Directed Behavior: Agents aren’t simply waiting for instructions, they’re designed to pursue defined goals, such as creating a summary report or onboarding a new user into a system.

Multi-Step Reasoning: Rather than stopping after one interaction, agents can plan, sequence, and iterate through a series of actions to reach an outcome.

Tool Use: Agents interact with other systems; APIs, search tools, internal databases, document repositories to retrieve, synthesize, and act.

This moves the role of system from responder to participant - it takes ownership of a task rather than waiting for a stream of commands.

Real-World Application

The potential of AI agents becomes especially clear in business processes that involve multiple steps, multiple tools working together, or repetitive effort. Generative AI agents don't just interpret or generate content, they run across knowledge sources, tools, and logic to complete tasks on their own. And these are not hypothetical scenarios - common day-to-day challenges that consume considerable time and cognitive effort from skilled professionals can be improved quickly. Simple examples to paint the picture:

Meeting summarization and follow-up: Rather than manually reviewing transcripts or notes, a generative AI agent could autonomously:

- **Summarize key decisions and discussion points.**
- **Extract and assign action items in a task management system.**
- **Draft and distribute recap emails.**
- **Schedule the next steps based on tasks or action items.**

Impact: According to Microsoft's 2023 Work Trend Index, inefficient meetings are the top productivity disruptor, with 56% of knowledge workers finding it difficult to summarize what happened in meetings (Microsoft, 2023). This could turn a 30-minute post-meeting admin task into an automated outcome that frees knowledge workers to focus on preparing for what comes next rather than recapping what just happened.

RFP response generation: In professional services and public sector work, responding to Requests for Proposals (RFPs) is time-sensitive and resource-intensive. A generative AI agent could:

- **Parse RFP requirements.**
- **Retrieve similar past submissions or boilerplate content.**
- **Draft tailored responses.**
- **Route the draft to right stakeholders for iteration and review.**

Impact: Reduces hours of manual document assembly and content retrieval allowing teams to focus on customizing strategy and messaging instead of formatting and searching.

Client or Employee Onboarding: Onboarding processes often require navigating multiple systems: HRIS, CRM, knowledge bases, scheduling platforms. A generative AI agent could:

- **Populate client or employee profiles based on intake forms.**
- **Initiate access requests and compliance workflows.**
- **Schedule orientation sessions and send communications.**
- **Generate FAQs and starter kits.**

Impact: Minimizes delays and inconsistencies by ensuring that repetitive onboarding steps are completed efficiently and in the right order—reducing reliance on overextended support teams

But despite the promise of AI agents, they aren't the right fit for every task. Simpler generative AI assistants are still ideal for direct, single-turn scenarios:

- **"What's our PTO policy?"**
- **"Summarize this article."**
- **"Translate this paragraph."**

Deploying an autonomous generative AI agent to answer a basic FAQ is overkill technically and operationally. Instead, generative AI agents should be reserved for use cases that benefit from multiple rounds of reasoning, autonomy, and orchestration across systems. Knowing the distinction helps organizations invest wisely and avoid complexity where it's not needed.

Governance, Oversight, and Risk

With autonomy comes responsibility. A recent analysis by Palo Alto Networks highlights that AI built on large language models (LLMs) can inherit security risks such as prompt injection, sensitive data leakage, and supply chain vulnerabilities (Palo Alto Networks Unit 42, 2025). In addition, these agents integrate external tools (per our definition of an Agent) developed in various programming languages and frameworks which may also expand the attack surface and potential for exploitation. When generative AI agents act on behalf of the organization, the risks shift:

- **Outputs aren't just suggestions, they're completed tasks.**
- **Mistakes aren't just model errors, they're operational events.**
- **Inferences can cascade into decisions, actions, or communications.**

The transition from reactive tools to active collaborators must be accompanied by thoughtful design, structured oversight, and active involvement from business and IT leaders. Generative AI agents demand stronger governance frameworks that should minimally include:

- **Clear boundaries on what actions agents can take.**
- **Human-in-the-loop controls for critical business processes.**
- **Auditability and explainability of actions taken.**
- **Security reviews for agent access to internal systems.**

Why it Matters for Knowledge Work

The rise of generative AI agents is a shift in the future of knowledge work. While early generative AI experiences have focused on generation and retrieval, delegated action will define the next wave. Generative AI agents will be empowered not just to help, but to collaborate and act on our behalf.

Understanding what makes a generative AI system an agent (and when an agent should be used) is becoming core to the Awareness pillar of generative AI literacy. As agents appear in more platforms and internal tools, knowledge workers should:

- **Recognize when they're interacting with an agent vs. an assistant.**
- **Understand the boundaries of agent autonomy.**
- **Know how to supervise, confirm, and escalate when needed.**

CONFIDENCE: BUILD TRUST THROUGH PRACTICAL EXPERIENCE

Awareness is the foundation of generative AI literacy, but awareness alone is not enough. For generative AI to meaningfully enhance knowledge work, knowledge workers must move beyond conceptual understanding and develop confidence using these tools. Confidence enables users to engage with generative AI systems not as passive recipients, but as active collaborators. They can ask better questions, challenge results, and set clear expectations about how generative AI contributes to their work.

But confidence is not about blind trust. It is built through practical experience, discussion, and the ability to provide feedback that the organization takes seriously. When knowledge workers interact regularly with generative AI, they begin to form the mental models that guide their responsible use of this technology. This confidence is what allows organizations to move from curiosity toward reliable integration.

In this section, we'll briefly explore how confidence grows through firsthand interaction, and an understanding of the real strengths and limitations. We'll also discuss how leaders can encourage an environment where exploration is encouraged, skepticism is healthy, and the goal is productive collaboration between human knowledge and generative AI systems.

PRACTICAL EXPOSURE DRIVES FAMILIARITY

Confidence in generative AI doesn't come from reading technical specifications or hearing success stories alone. It comes from direct engagement. Just as professionals build intuition for a new software platform by using it day to day, generative AI tools must become part of routine business processes before users can fully understand their proper role.

Structured exposure through pilots, guided experiments, or sandbox environments gives knowledge workers a safe space to explore how these systems behave in the real world. Accenture's research indicates that organizations investing in AI literacy programs see a 35% improvement in productivity, underscoring the value of firsthand experience (Accenture, 2023).

For instance, a policy analyst might use a generative AI tool to summarize a new regulation, then compare the output to their own interpretation. Or a customer support lead might generate a draft response to a client issue and then refine it manually. These moments of side-by-side evaluation help users build trust and improve their own methods of interaction.

Familiarity breeds not only better outputs but better questions. As users interact with these tools across a range of scenarios, they begin to recognize patterns: where AI tends to oversimplify, where it performs well, and what types of inputs yield the most useful results. This learning loop builds a form of human-AI fluency that is hard to replicate through theory alone.

It's important to design these first-hand opportunities intentionally. Not every use case is ready for generative AI integration. But use cases that are ready for generative AI should serve as internal case studies that show real value while giving users the confidence to think more broadly about how AI can support their work.

RECOGNIZING STRENGTHS AND LIMITATIONS

Confidence in generative AI is about developing "informed trust." This means understanding not just where generative AI shines, but also where it struggles, and why.

Generative AI systems excel at summarization, rewriting, brainstorming, translating, and adapting information to different tones or audiences. They are particularly useful when the input data is rich in structure and context. In these conditions, generative AI can surface insights, draft responses, and reduce the time spent performing repetitive tasks. Examples of context-rich data could include company policy documents, project histories, or customer communications.

But confidence also comes from recognizing generative AI's current limits. It is prone to plausible-sounding but incorrect responses - "hallucination." And just like humans, it may misinterpret vague questions, struggle with outdated or domain-specific knowledge, or apply formatting incorrectly in complex document structures. But unlike human counterparts, generative AI will try to give an answer. And in a strange way this is why generative AI hallucination is such an important problem to understand. An incorrect response confidently stated will often be accepted as fact.

Equipping knowledge workers with this balanced perspective, where strengths and limitations are well-understood should encourage them to engage more critically. They can better evaluate results and confirm these results against trusted sources rather than relying on them blindly.

Confidence, then, is not simply about believing the system works. It's about knowing when it works, why it works, how to confirm, and what to do when something isn't right.

CREATING A CULTURE OF TRANSPARENCY AND DIALOGUE

Building confidence in generative AI is an ongoing cultural practice. Confidence grows when organizations create space for curiosity, transparency, and open conversation about how these systems work, where they're being used, and what outcomes they're driving.

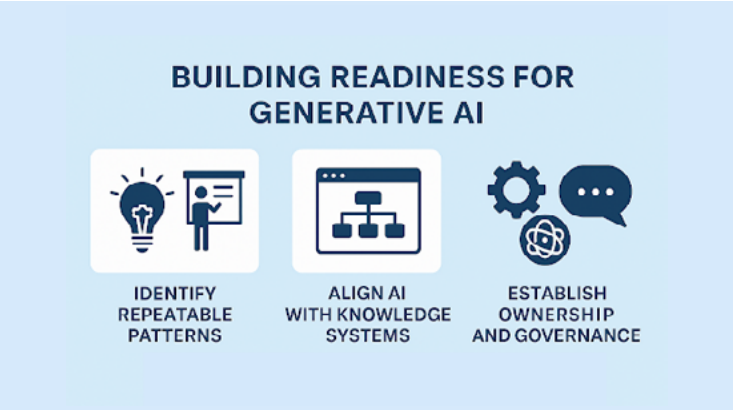
A study conducted by KPMG in collaboration with the University of Melbourne found that of the 48,340 individuals across forty-seven countries surveyed, 57% of employees admitted to hiding their use of AI at work (KPMG & University of Melbourne, 2025). Too often, generative AI systems are introduced behind the scenes, with limited explanation or visibility. When employees don't know how a system generates its output, what data it's drawing from, what data protection policies should be in place, or how to flag errors trust and confidence erodes.

Instead, confidence is built through clarity. This includes explaining which tools are available for use and why, what data sources are involved, and what guardrails are in place to mitigate the risks for the organization and the individual. When leaders actively communicate the *why* and *how* behind AI initiatives, they invite engagement, not resistance or shadow AI initiatives.

Creating this culture also means encouraging feedback. Knowledge workers should feel empowered to question results, advocate for better tools, escalate when something doesn't look right, and collaborate with technical teams to refine use cases. Healthy generative AI implementation happens when the systems and its users are both learning. And, where conversations about limitations are valued alongside celebrations of success.

Transparency is a confidence accelerator. And in an environment where knowledge workers feel safe to explore, critique, and improve how AI is used, confidence becomes a shared asset across the organization.

READINESS: OPERATIONALIZING GENERATIVE AI IN THE FLOW OF WORK



Many organizations begin their generative AI journey with great enthusiasm, enticed by the potential generative AI offers to improve the quality and efficiency of their work. They experiment in sandboxes, run pilots, and explore a variety of tools. But without a clear path to more meaningful integration, these experiments often do not scale. That's where readiness comes in. Readiness is the bridge between isolated innovation and broader transformation with generative AI. It is the ability to embed generative AI tools into daily business processes in a way that is intentional and has potential for real impact.

This pillar of Generative AI Literacy shifts the focus from conceptual awareness and early confidence-building toward practical implementation. It's not just about building or adopting new tools, it's about preparing people, processes, and policies. Readiness ensures that generative AI doesn't remain a novelty delegated to a small team of early adopters. Instead, it can become a core capability that supports important knowledge processes and empowers frontline decision-making.

WHY READINESS MATTERS

Generative AI offers exciting potential to transform the way knowledge work is performed. It can automate repetitive tasks, synthesize insights from enormous amounts of unstructured content, generate high-quality drafts in seconds, and personalize knowledge delivery at scale.

But realizing that potential depends on how deeply it is woven into the operational fabric of the organization. Without readiness, generative AI projects can remain as experiments or in the shadows. They are interesting and potentially useful to the individual but relegated to the periphery. In fact, McKinsey reports that only 11% of companies have successfully scaled generative AI (McKinsey & Company, 2024).

Some organizations that were early generative AI adopters face the common challenge of isolated use cases that don't scale. A team may pilot a generative AI tool to summarize reports or help with documentation, but without a structured path forward (governance, integration, support) this promising pilot stays siloed. These "innovation islands" are symptoms of a deeper issue: a lack of operational readiness to absorb generative AI into the way work gets done.

The readiness pillar addresses this by helping ensure that generative AI is not introduced as an extra layer, but as a meaningful business capability. That means aligning tools with specific business processes, training teams to use them responsibly, and setting up feedback mechanisms to continuously improve how generative AI supports knowledge work. It also means working across business units, IT, and compliance teams to create conditions where experimentation can safely scale into practice. Readiness is what converts potential into performance.

KEY ENABLERS OF GENERATIVE AI READINESS

Building readiness for generative AI integration isn't only about deploying technology. It's about aligning the right people, processes, and infrastructure to support it. The following enablers can help organizations transition from isolated experimentation to enterprise-wide adoption:

Executive Sponsorship with a Clear Use Case Strategy: Leadership buy-in is essential, but it must be paired with clarity. What are the specific knowledge-intensive processes generative AI will improve? Contract drafting, internal Q&A, policy summarization? Well-defined use cases help teams prioritize where to focus and how to measure success.

Structured Experimentation Environments: Sandboxes and pilot programs provide a safe space for testing generative AI tools without disrupting core systems. These environments should allow users to explore capabilities, evaluate real-world prompts, and find integration needs while capturing feedback to inform broader rollouts.

Cross-Functional Collaboration: Embedding generative AI into knowledge workflows requires alignment across teams. Knowledge workers, IT, data stewards, compliance, and security. Use cases than span departments can encourage shared buy-in. Readiness increases when these groups work together early to design solutions that are practical, secure, and include the right amount of context.

Data and Access Governance: Generative AI thrives on access to high-quality content, but it also raises questions of data sensitivity. Organizations must define governance guardrails that ensure the right data is accessible to AI systems without introducing risk.

Workforce Enablement and Training: Even the best tools underperform without empowered users. Organizations should offer role-based training, prompt design best practices, and real-world examples to help teams use AI confidently and responsibly. This complements the "Confidence" pillar and reinforces a feedback loop between use and learning.

Continuous Feedback and Optimization Loops: Organizations should put mechanisms in place to continuously learn from how generative AI is being used. This feedback should drive improvements to both the tools and the supporting infrastructure.

MEASURING PROGRESS

Organizational readiness isn't a binary achievement, it evolves over time. To ensure that generative AI moves from experimentation to meaningful adoption, leaders need a way to measure progress. While traditional IT metrics focus on uptime, deployment frequency, and time to value; generative AI readiness requires a different lens. One that blends user engagement, business process integration, and strategic alignment.

Use Case Maturity: Track the lifecycle of generative AI initiatives across the organization.

- **Exploration:** Are teams actively discovering and proposing use cases?
- **Piloting:** Are experiments running with clear goals and evaluation criteria?
- **Scaling:** Are successful pilots expanding into departments or enterprise-wide programs?

Business Process Integration Rate: Evaluate how well AI capabilities are embedded into daily processes. Are generative AI tools part of how teams author reports, summarize meetings, or respond to RFPs? Look beyond tool usage stats to measure how often AI outputs are informing real decisions and deliverables.

Adoption and Confidence Metrics: Use engagement metrics like active users, repeat usage, and prompt complexity to assess how comfortable users are becoming. Pair this with surveys or feedback loops that track perceived trust, value, and clarity around responsible use.

Collaboration and Support Readiness: Assess whether core teams are working together effectively. IT, data governance, security, business units. Are shared policies in place for content validation, prompt libraries, or data access? Are issues being surfaced and resolved collaboratively?

Time to Value: Monitor how long it takes for generative AI initiatives to move from idea to impact. Are barriers slowing adoption? Is data access limited, or is ownership unclear? Reducing time-to-value is a key indicator of increasing readiness.

Feedback Loop Strength: Track the presence and responsiveness of feedback mechanisms. Are teams reporting successes and failures? Are updates being made to prompt libraries, data strategies, or guidance materials based on real use?

FROM PILOTS TO PRACTICE

Pilot projects are valuable, but they are only the beginning. For generative AI to drive long-term impact, organizations must shift from isolated experimentation to sustained operational integration. That means embedding generative AI into the rhythm of knowledge work. Not as a novelty, but as a reliable tool that complements and extends human ability.

Find Repeatable Patterns: Successful pilots often uncover workflows that are ripe for repeatability. Summarizing client meetings, drafting internal memos, or extracting key insights from policy documents. These are prime candidates for inclusion into a use case. Once a pattern is recognized, create reusable prompt libraries, data access templates, and user guidelines to accelerate expansion.

Build Enablement Pathways: Organizations need to move beyond pockets of ability and create role-specific enablement pathways. Train HR professionals on AI-driven policy, enable legal teams to confirm generative AI-assisted document drafts, and encourage customer support teams to help tune prompts. Readiness scales when learning is role-aware and context-specific.

Align AI with Core Knowledge Systems: For generative AI to move into practice, it must connect with the systems and sources knowledge workers already rely on. Content management systems, document repositories, ticketing platforms, and communication tools. RAG-based systems require integration with curated knowledge stores to ensure responses are grounded, secure, and aligned with internal truth.

Establish Ownership and Governance: Sustained success requires more than technical readiness; it requires clear ownership. Who updates prompt libraries? Who confirms system outputs in regulated environments? Who watches for drift in AI responses over time? Assigning roles and responsibilities ensures that generative AI doesn't remain a shadow tool but becomes part of the organization's operating fabric.

Normalize AI Collaboration: Most importantly, organizations must help teams normalize collaboration with generative AI tools. This means recognizing that AI doesn't replace knowledge workers. In high-performing environments, AI becomes a second brain: drafting, exploring, retrieving, and organizing on demand while humans apply context, judgment, and creativity.

CONCLUSION

Generative AI represents more than a technological shift. It signals a fundamental transformation in the landscape of knowledge work. Across industries, organizations are already tapping into generative AI's potential by unlocking efficiency, building new capabilities, and deriving deeper insights from their vast stores of information. However, capturing sustained value from this powerful technology demands more than tool adoption. It demands widespread generative AI literacy.

As outlined throughout this paper, building this literacy requires deliberate investment across three interdependent pillars:

- **Awareness:** Equipping your workforce with the foundational understanding of generative AI's core concepts, including the critical distinctions between training and inference, and discriminative versus generative AI. This foundation empowers knowledge workers to use these tools confidently and accurately.
- **Confidence:** Establishing practical trust through firsthand exposure, transparent communication, and clear expectations around generative AI's capabilities and limitations. Organizations that cultivate user confidence are better positioned to drive adoption and avoid common pitfalls associated with opaque AI deployments.
- **Readiness:** Preparing your organization's culture, processes, and governance structures to fully integrate generative AI within business processes. Readiness ensures AI technologies move beyond isolated experimentation into robust, repeatable, and scalable capabilities embedded within your organization's operational fabric.

Ultimately, the organizations that win in this new era won't simply be those with access to advanced technology. Instead, they'll be the ones whose people deeply understand how to harness generative AI to enhance - not replace - their domain expertise and professional judgment. By investing deliberately in generative AI literacy, organizations can empower their workforce, unlock tangible productivity gains, and build sustainable competitive advantage.

Generative AI literacy is no longer optional. It is the new imperative for technology leaders. Now is the time to define your organization's roadmap clearly, identify meaningful use cases, and begin the practical journey towards widespread AI literacy. Doing so positions your organization not just to adapt to the future of knowledge work - but to define and lead it.

YOUR CLOUD JOURNEY STARTS HERE

Whether you're modernizing infrastructure or optimizing operations, AIS can help you turn today's needs into tomorrow's outcomes. Let's build a roadmap that puts transformation into action—with clarity, speed, and purpose.

References

1. Mattioli, D. (2024, May 5). Turns Out AI Is More Empathetic Than Allstate's Insurance Reps. The Wall Street Journal.

   https://www.wsj.com/articles/turns-out-ai-is-more-empathetic-than-allstates-insurance-reps-cf5f7c98

2. ProProfs Knowledge Base. (2023, October 5). Workforce spends too much time searching for information: Study finds. ProProfs.

   https://www.proprofskb.com/blog/workforce-spend-much-time-searching-information/

3. KPMG & University of Melbourne. (2025). Trust in artificial intelligence: Global insights 2025. KPMG Australia.

   https://kpmg.com/au/en/home/insights/2025/04/trust-in-ai-global-insights-2025.html

4. Microsoft. (2023). 2023 Work Trend Index Annual Report: Will AI Fix Work?. Microsoft Corporation.

   https://assets.ctfassets.net/y8fb0rhks3b3/5eyZc6gDu1bzftdY6w3ZVV/beeb0f2e437044fc99a0408867a263d8/WTI_Annual_2023_Will_AI_Fix_Work_.pdf

5. Accenture. (2023). The productivity payoff: AI and the future of work. Accenture.

   https://www.accenture.com/content/dam/accenture/final/accenture-com/document-3/Accenture-The-Productivity-Payoff.pdf

6. LinkedIn. (2024). Work Change Report: Navigating the new world of work. LinkedIn Economic Graph.

   https://economicgraph.linkedin.com/content/dam/me/economicgraph/en-us/PDF/Work-Change-Report.pdf

7. McKinsey & Company. (2024). From promising to productive: Real results from gen AI in services.

   https://www.mckinsey.com/capabilities/operations/our-insights/from-promising-to-productive-real-results-from-gen-ai-in-services

8. Palo Alto Networks Unit 42. (2025). AI agents are here. So are the threats. Palo Alto Networks.

   https://unit42.paloaltonetworks.com/agentic-ai-threats/

9. OpenAI. (2024). GPT-4o: OpenAI's latest multimodal model.

   https://openai.com/index/gpt-4o

ais

www.ais.com

11440 Commerce Park Drive, Suite 600, Reston, VA 20191

Phone: (703) 860-7800 Fax: (703) 860-7820

sales@ais.com