Market Momentum Index:

AI and Unstructured Data Management

Research by Deep Analysis in partnership with AIIM

Deep Analysis

aiim
Association for Intelligent Information Management

Underwritten by

M-Files

Foreword

At AIIM, we believe unstructured data and information management is intrinsically linked to success with Artificail Intelligence (AI). As fuel for model training and prompts, unstructured data is essential for success with AI. Information management is the practice that curates, manages, and protects unstructured data and ensures its usability for AI.

However, all of that is belief.

As an industry, we lack clear understanding of how AI will shape the unstructured data management space and the speed of that change. We lack definitive data that justifies our belief that information management is a valuable and essential practice in the AI era.

This important research by Deep Analysis, supported by M-Files answers three critical questions:

1. *How fast are organizations adopting AI? What are the barriers to adoption?*
2. *What is the state of organizational readiness for AI?*
3. *What is the level of importance of unstructured data in the implementation of AI?*

In other words, this research examines the very relevancy and future of the information management industry.

What you will discover in this gutsy report is that we are adopting AI faster than we previously thought. 77% of survey respondents said that their organizations had AI projects in evaluation or production.

Most organization feel ready for AI and have taken important steps towards readiness. The data also shows that unstructured data plays an important role in this paradigm and is also going to be impacted by AI. AI will change the way we manage unstructured data going forward.

Contrary to the belief that organizations are worried about the dangers of AI to the point of immobility, this research shows that organizations are eager to embrace AI capabilities and users are ready. Stakeholder adoption was not a concern for respondents.

Organizations are also taking a pragmatic approach to AI. For example, survey respondents were more concerned about the security and performance of AI than they are about AI hallucinations despite the focus on AI mistakes in the media. And while we might be craving more publicly available use cases, 92% of organizations have identified processes that could be improved with AI.

The research demonstrates that this pragmatic approach to AI is enabled by the "free, fresh flow" of unstructured data.

Survey results show organizations are planning on increased investment in unstructured data management. While compliance may still be a primary motivator for investing in information management, it's clear that the times are changing.

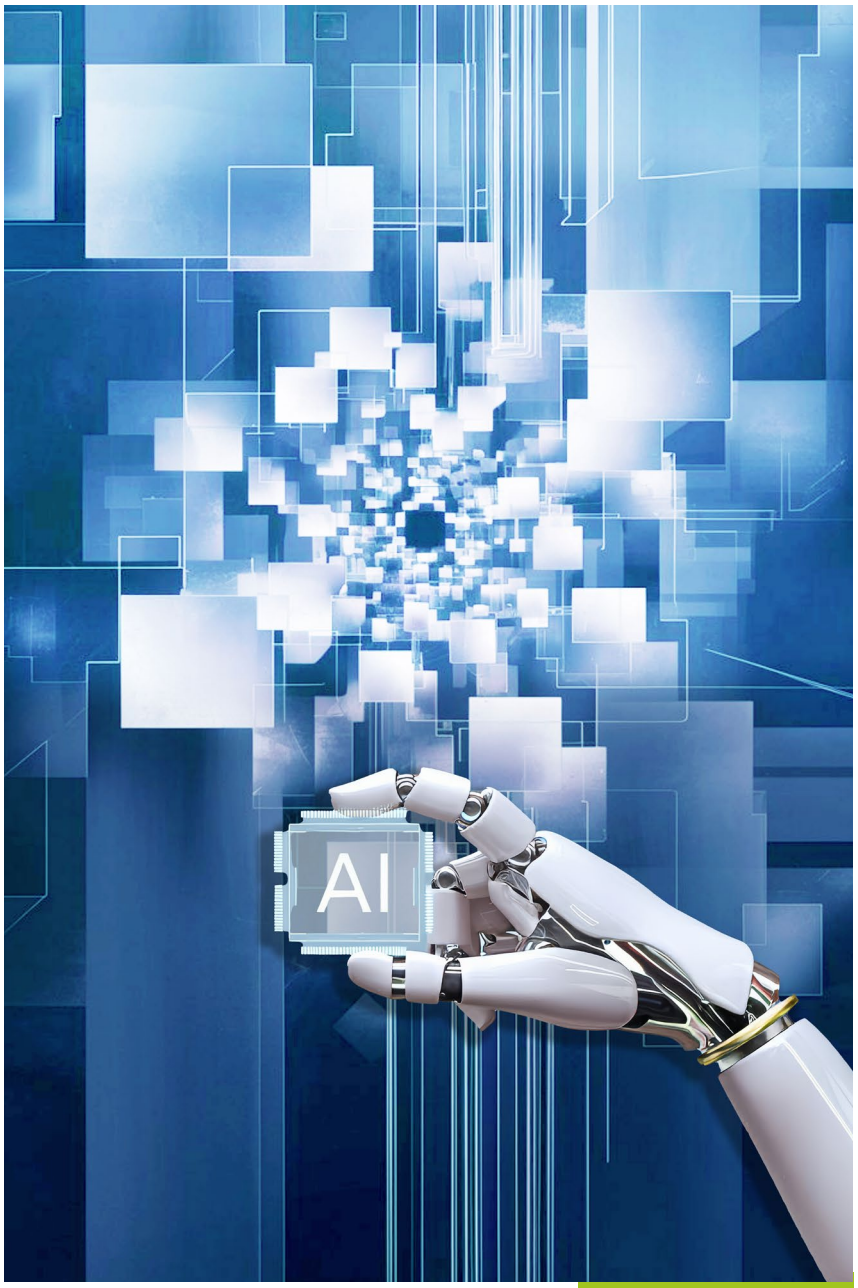
For success with AI, the report shows the importance of managing unstructured data across a growing multitude of systems; the need for enterprise automation and a holistic, end-to-end approach to workflow automation; the need for process mining to surface opportunities for AI improvements; and the still-prevalent need to leverage (and potentially) transform paper-based process.

The well-established practice of information management is the answer to the new needs of the AI era.

AI and unstructured data management aren't just partners—they're a powerhouse duo. This research isn't merely informative; it's a call to action. Use these insights to transform your organization's approach, turning data complexity into AI-driven advantage. The future favors those who excel at both.

aiim
Association for Intelligent Information Management

Tori Miller Liu, MBA, FASAE, CAE, CIP
President & CEO
Association for Intelligent Information Management (AIIM)



Introduction

It has been impossible to ignore the recent surge of artificial intelligence (AI) to the top of every business-focused technology provider's agenda. Whether AI is the intended primary topic of discussion or not, it inevitably surfaces, with the underlying message that neglecting AI is akin to ignoring an essential and imminent advantage that every organization should urgently seize.

However, to truly harness the transformative power of AI, organizations must prioritize robust information management. AI's potential is deeply intertwined with the quality and accessibility of the data it processes—especially the vast amounts of unstructured data where much of an enterprise's knowledge resides. Without a strategic approach to managing this information, the benefits of AI may remain out of reach.

As organizations increasingly turn to AI to tackle their unique business challenges, the need for effective information management becomes ever more critical. The processes and tasks that define how organizations operate present a substantial opportunity for AI to leverage existing enterprise knowledge. This, however, requires a foundation of well-organized, secure and accessible data that AI can use to enhance operational efficiency and drive innovation.

In this context, AIIM and M-Files commissioned Deep Analysis to conduct a research project to explore the role of unstructured data and the software managing it in the enterprise adoption of AI. The primary goal was to assess the readiness of organizations to employ AI, specifically focusing on how unstructured data is being prepared for AI utilization and identifying the tools, techniques, and skill sets that are supporting these efforts. Understanding these elements will enable organizations to better plan and execute their AI strategies, armed with insights into how their peers are navigating similar challenges.

> *"To navigate the future of work, you must harness the power of information to enable knowledge workers to thrive. If you have the right information management foundation, AI and automation can help enterprises understand the context of documents and interact with their knowledge using natural language accurately and in confidence to gain competitive advantage."*

For this project, 500 enterprises were surveyed through a questionnaire. All participants are U.S.-based, with revenues exceeding $1 million, and represent a cross-section of key industries.

Quantitative research data was collected during June and July 2024 and analyzed in July and August 2024. Full details of the data collection, sampling criteria, and other associated information can be found in the methodology section of this report.

Executive Summary and Key Findings

The first *Market Momentum Index: AI and Unstructured Data Management,* conducted by Deep Analysis with support from AIIM and M-Files, reveals that AI is already far more embedded into the operations of organizations than was previously realized.

Of the 500 US-based enterprises surveyed by Deep Analysis, **77% have AI projects either in production or in evaluation, and of that group, 79% have projects in production.** The enterprises are in the following sectors: Financial and Insurance; Professional, Scientific and Technical Services; Manufacturing; Mining, Quarrying, and Oil and Gas Extraction; and Utilities.

Though AI is not new, Generative AI is relatively new and has been hogging the headlines over the past 18 months. Prior to the launch of ChatGPT and the like, AI was considered something of a niche product in the sector, one that was used sparingly. This survey indicates that it is far more broadly in use than many realize; in short, buyers are not as green or uninformed regarding AI as is commonly thought.

The survey revealed a number of additional key findings that are summarized here.

- Looking toward specific readiness to implement AI, 92% of respondents said their organizations have identified at least one existing process that they believe will be improved with the introduction of AI. Of those respondents, 87% identified specific tasks within those processes, and 63% identified multiple such tasks.

- The survey suggests strong agreement that AI will impact how unstructured data is managed going forward. Organizations report that they have located the knowledge bases of information to support their AI projects and are confident in their quality and accuracy.

- Companies in Professional, Scientific and Technical Services (20%) and Financial and Insurance (19%) identify themselves most commonly as using AI in a "transformational" way (compared with a mean of 14% across the entire sample). However, the percentage across all analyzed verticals in terms of depth of evaluation and production trends is nearly uniform, which suggests that in terms of adoption, applications for horizontal business functions are currently running ahead of those that apply to individual industries themselves.

- Despite commonly reported concerns about AI "hallucinations," survey respondents cited security and access compliance (43%) and performance and accuracy (40%) as the biggest factors preventing enterprises from effectively leveraging AI. AI hallucinations (31%) and ethical and bias concerns (25%) were raised less frequently. This provides a strong indication that organizations are focusing on the practical factors on which the success of their AI deployments rest, rather than the more headline-grabbing hyperbole that has accompanied the rise to prominence of the technology itself.

- Those knowledge bases of information are contained in a wide range of repositories, with respondents typically saying that they are spread across multiple applications, each containing no more than one-fifth to one-quarter of the total information.

- Organizations said that typically 67% of their unstructured data resides in cloud storage, accounting for 57% of their overall storage volume. Only 4% of those organizations are seeing any decrease in costs of managing that unstructured data, while 55% said their associated costs are increasing.

- Software (50%) rather than storage (31%) accounts for most of that cost increase for managing unstructured data, with the remainder made up by spending on employees and external skills. Enterprise automation (69%) and knowledge management (61%) software are the two largest categories within that increased software spend. RPA and low/no code systems are the lowest priorities (both at 24%).

Methodology

Quantitative data for the *Market Momentum Index: AI and Unstructured Data Management* survey was gathered via a set of panel questionnaires conducted anonymously within 500 US-based enterprises.

The quantitative data panel was created from self-identified respondents matching the following profile, with no weightings unless specified:

Geographical Profile: Based in the United States

Industry Profile (equally weighted within sample):

- Financial and Insurance
- Professional, Scientific and Technical Services
- Manufacturing
- Mining, Quarrying, and Oil & Gas Extraction
- Utilities

Financial Profile: annual revenue greater than $1 million

Size Profile: More than 500 employees

Employment Role:

- Executive/C-level (CEO, CIO, CTO, CDO, etc.)
- VP level
- Director level
- Manager level

Organizational Structure:

- IT Organization
- Digital Transformation/Innovation Office
- R&D
- Product Development/Engineering
- Manufacturing/Operations
- Corporate Strategy & Planning

Analysis of Research Data

This inaugural version of the *Market Momentum Index: AI and Unstructured Data Management* survey, conducted in the summer of 2024, was conceived with the support of AIIM and M-Files to reveal the degree to which organizations are adopting artificial intelligence (AI), the extent of organizational readiness for that adoption, and the role unstructured data is playing in that dynamic.

As discussed in the Executive Summary/Key Findings, many areas of interest were uncovered during the quantitative research that help illustrate the current situation within the selected sample of organizations, verticals, and roles/departments. In this section, we expand upon those areas of interest, with selected illustrative charts providing additional insights.

Operational, accelerated, transformational: current use of AI

One of the most striking findings from this survey is the revelation that 77% of the responding organizations have AI projects either in production or in evaluation. Breaking it down, 61% of all respondents have AI projects in production (operational, accelerated, and transformational [OAT] responses), while 16% are experimenting *(see Figure 1)*. Perhaps the biggest takeaway is that the overwhelming majority of organizations that have taken action on AI (79% of this cohort) have passed the experimentation phase and deployed AI into their day-to-day operations.

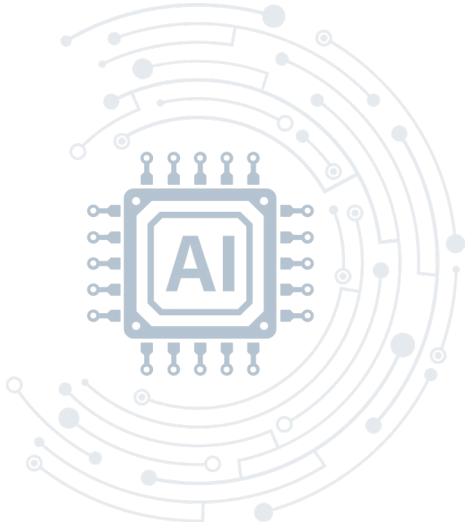
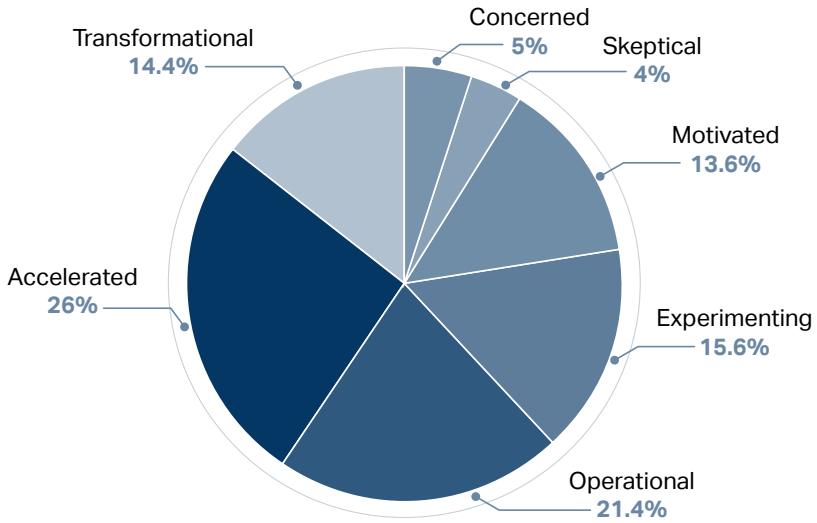


Figure 1 - AI Projects in Production or in Evaluation

[S7] As it relates to the use of AI (including GenAI and/or LLMs) within business, at what stage would you place your organization? (n=500)



Within that cohort, the results are even more illuminating. The largest response is for "accelerated," suggesting that it's far from just special edge cases being deployed. Over one-quarter of the entire survey sample are looking to deploy AI across the board wherever they spot an opportunity.

Looking at the industry verticals and the Operational, Accelerated, Transformational cohort of responses, it is clear that AI is being actively deployed across all five chosen industries and on a largely consistent basis, albeit with some differences between the specified phases of operation *(see Figure 2)*. The combined production bar in the chart aggregates the total percentage of respondents across all three phases—Operational, Accelerated, and Transformational—providing an overall view of AI adoption in production. Respondents in the Financial and Insurance and Professional, Scientific and Technical Services verticals are most likely to see their operational approach as transformational *(for the detailed definition of this term and others used here, see the appendix reference for question S7).*

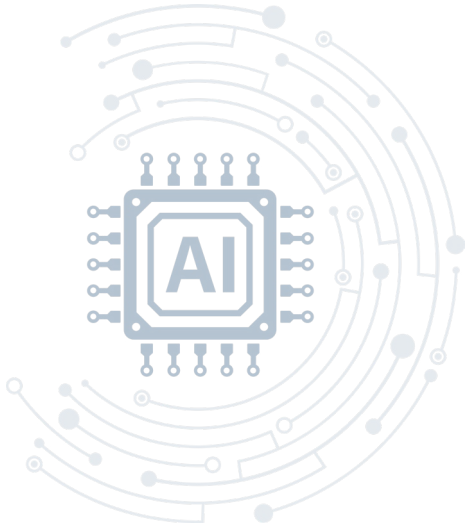
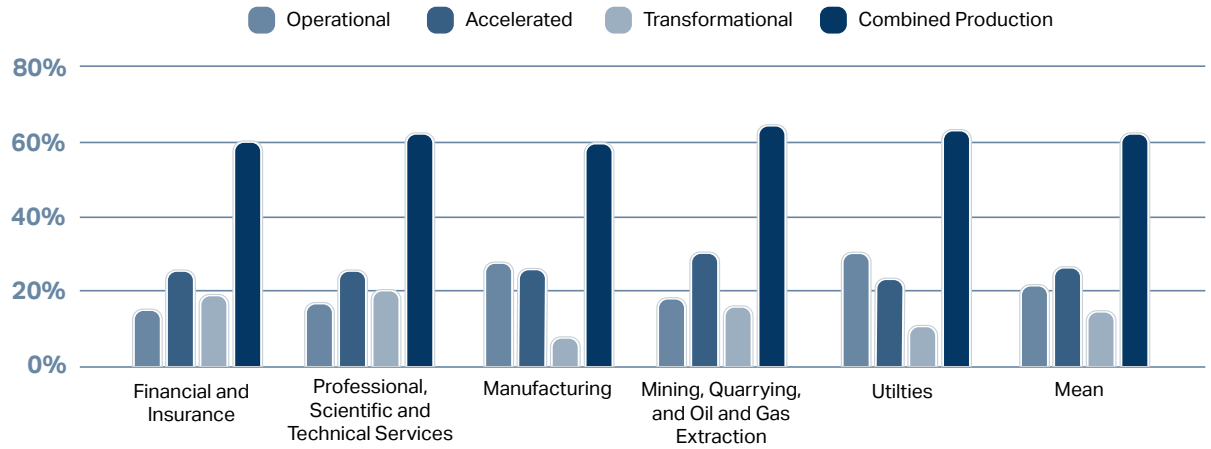


Figure 2 - AI Projects in Production or in Evaluation: By Vertical and Operational Projects

[S7] "As it relates to the use of AI (including GenAI and/or LLMs) within business, at what stage would you place your organization?" Broken down by vertical and operational projects? (n=500)



Mining, Quarrying and Oil & Gas Extraction has a marginally larger overall percentage of respondents in production than the other industries. With operational the lowest phase of actual use and transformational the most advanced, we can say that Manufacturing and Utilities lag slightly at 27% and 30% operational, and 7% and 10% transformational, respectively). The differences, however, are not distinct enough to draw direct conclusions, other than to wonder if what are being deployed are largely horizontal and common in nature (e.g., standard desktop, productivity functionality) rather than specific industry solutions.

It's important to note that each OAT label has a significantly subjective element. The labels perhaps best represent respondents' attitudinal state as much as their operational state. Therefore, it's best to interpret these labels as indicative of their organizational drive rather than truly representing the complexity of their actual AI deployments.

Accuracy and control rather than hyperbole

Much has been written since the initial market entry of generative AI (GenAI), both boosterish and pessimistic. Two of the most cited negative characteristics have been "hallucinations" and model bias and associated ethics.

"Hallucinations" – where a GenAI model returns incorrect information that it has synthesized while attempting to produce a complete response – have attracted the greatest volume of hyperbole; the term suggests that GenAI is a consciousness that can perform outside of human control, when in reality it is doing nothing of the sort. It is simply making incorrect guesses in attempting to fulfill the instruction ("prompt") it is processing. Differently, worries about model bias (referring to the selection of the information being supplied to create the models) and the ethics around model creation and management are no phantom; there are genuine concerns about the use of the large-scale, public models (LLMs/foundation models) whose users are typically blind to the source materials used to construct them (and any weighting of those sources within the models themselves).

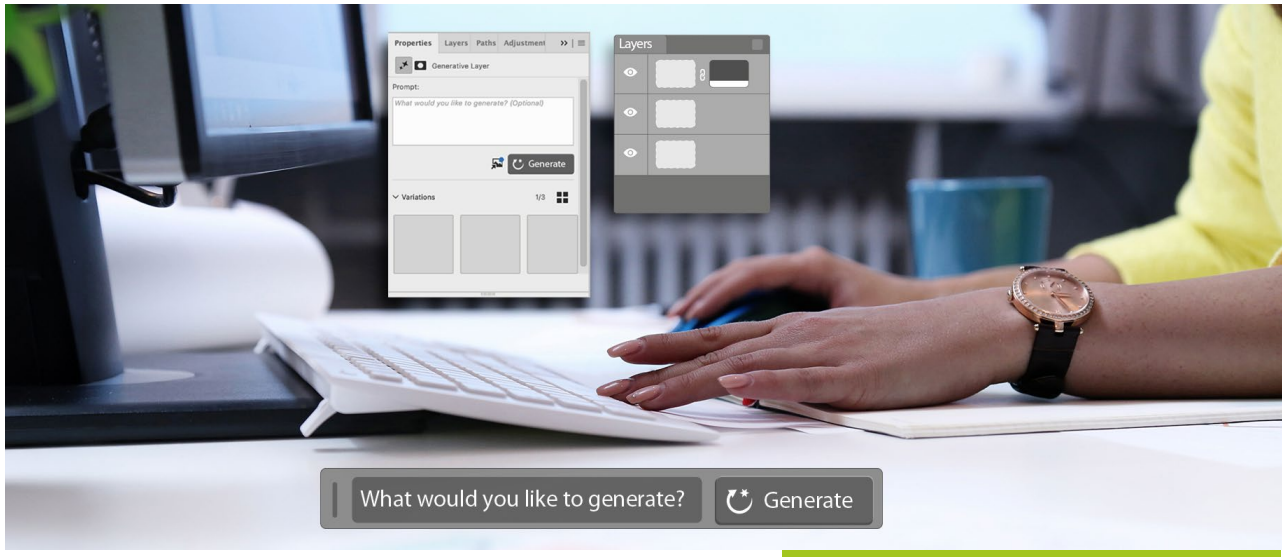
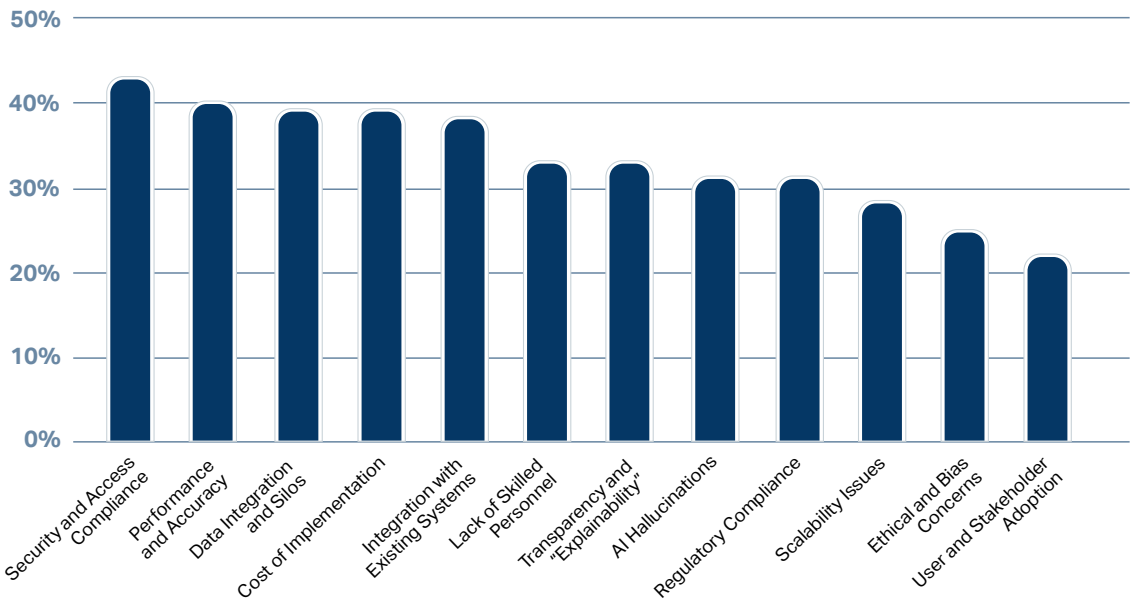
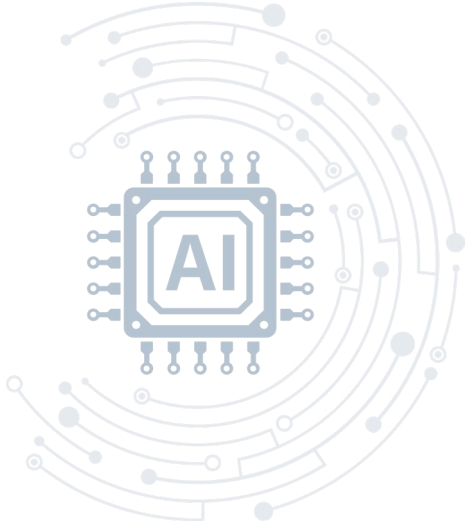


Figure 3 - Concerns Around Adopting AI

[Q28] "What prevents your enterprise from effectively leveraging AI?" (n=500)





In response to the question "What prevents your enterprise from effectively leveraging AI?" with multiple answers permitted, 31% of respondents selected AI hallucinations and 25% selected ethical and bias concerns *(see Figure 3)*. The most frequent responses were not those headline-grabbers, but more ordinary concerns: "security and access compliance" and "performance and accuracy."

This may be because the respondents are already experienced at bringing AI applications into production. They likely have experience with the complexity of fitting AI functionality within existing security and audit controls and making sure that it meets compliance standards. Additionally, they likely have conducted user acceptance testing (UAT), where accuracy and performance characteristics come under significant scrutiny.

That three IT-heavy responses come next – data integration, cost of implementation, and systems integration – is also instructive as far as lessons learned and concerns up front. Organizations that are embarking on integrating existing knowledge systems into their AI applications, for example using retrieval augmented generation (RAG), will find these very quickly on their whiteboards.

Reassuringly – given the AI production experience that our sample suggests – the lowest-scoring concern was around user adoption, suggesting that resistance from stakeholders and users in general has not been an overwhelming issue.

Processes and tasks – indicators of AI readiness

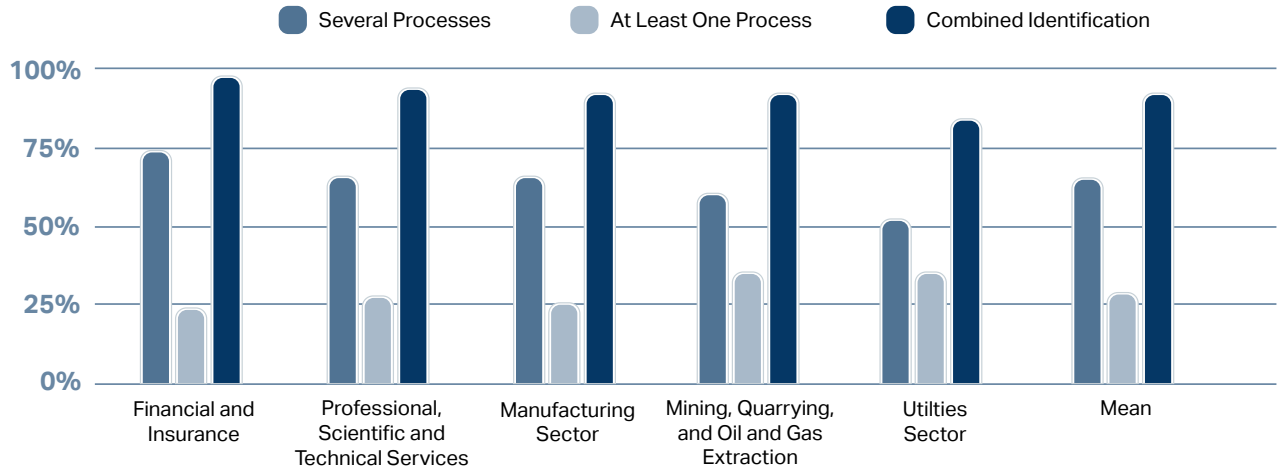
Prior to beginning the research project that created ***Market Momentum Index: AI and Unstructured Data Management*** in mid-2024, Deep Analysis and AIIM published an infographic (https://info.aiim.org/is-your-organization-ready-for-ai) called "5 Questions – Is Your Organization Ready For Artificial Intelligence?" The first of those questions is "Have you identified the processes that you think AI will improve?" Including this question in this survey allowed us to see whether organizations are taking that important first step.

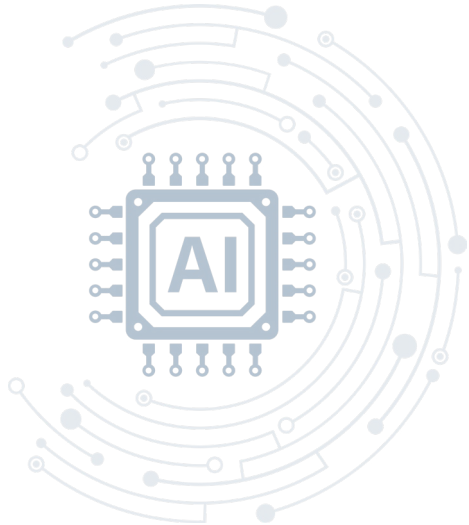
The results are extremely positive *(see Figure 4)*. Combining the responses for "at least one" with "several," 92% of respondents have identified processes that they believe will be improved by introducing AI, with 63% having identified multiple processes. Keeping in mind respondents' experience in putting AI applications into production, this points to a laudable approach toward process identification for AI.

The larger vertical breakout in responses shows that at the high end, 97% of Financial and Insurance respondents have identified processes (74% for several, 23% for at least one), while at the low end, 85% of Utilities respondents have done so (51% for several, 34% for at least one).

Figure 4 - Respondents Who Have Identified Processes They Believe AI Can Improve

[Q18)]"Have you identified processes in your organization that you think AI will help improve?" (n=500)





When asked subsequently, "Are you planning to use any specific process identification software tools (Process Mining) to assist you?" [Q22], 85% replied positively, with 44% adding that they are specifically using process mining software. This suggests a strong degree of professional process analysis within our sample; this is a positive sign for the likely success of the AI projects being launched.

The second question on the infographic was also asked in the survey [Q24], "Do you know the specific tasks within those processes that AI will power?" Here, 87% responded positively (combining "a number of" and "at least one"), with 63% having identified multiple tasks.

Unstructured data: vital, disparate, and challenging

Respondents broadly agree that AI will impact how unstructured data is managed within their organizations (90% responded yes [Q17]) and that its use (specifically "internal sources of knowledge") is important to the success of their organization's AI projects (80% median response [Q26]).

As organizations plan for or implement internal, unstructured data-heavy RAG systems – whether explicitly through software development or implicitly through packaged management like Microsoft Graph – the veracity of these knowledge sources will fundamentally impact how well the AI applications built on top of them function. Remember, too, that performance and accuracy was one of the top-cited concerns.

Survey respondents provided interesting insights regarding two elements: identifying sources of knowledge (Figures 5, 6, 7) and validating their quality (Figure 8). Regarding identification, 89% of respondents say they have identified both single and multiple sources, with 58% having located multiple sources and 31% one source *(see Figure 5)*.

> *"Companies need to break down information silos so that employees can leverage information from multiple structured and unstructured sources providing a 360-degree view of information and giving them the most accurate and up-to-date data for improved decision-making. Until that problem is solved, AI will not return the expected results."*

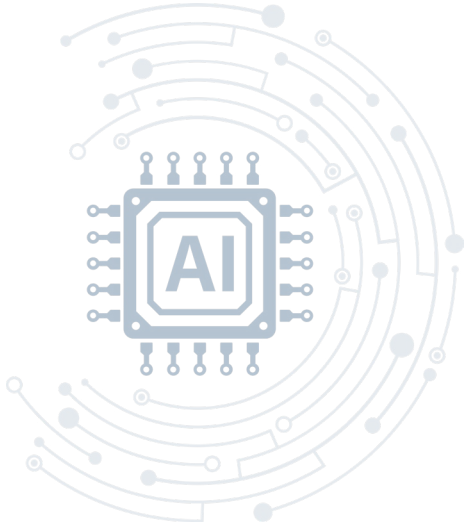
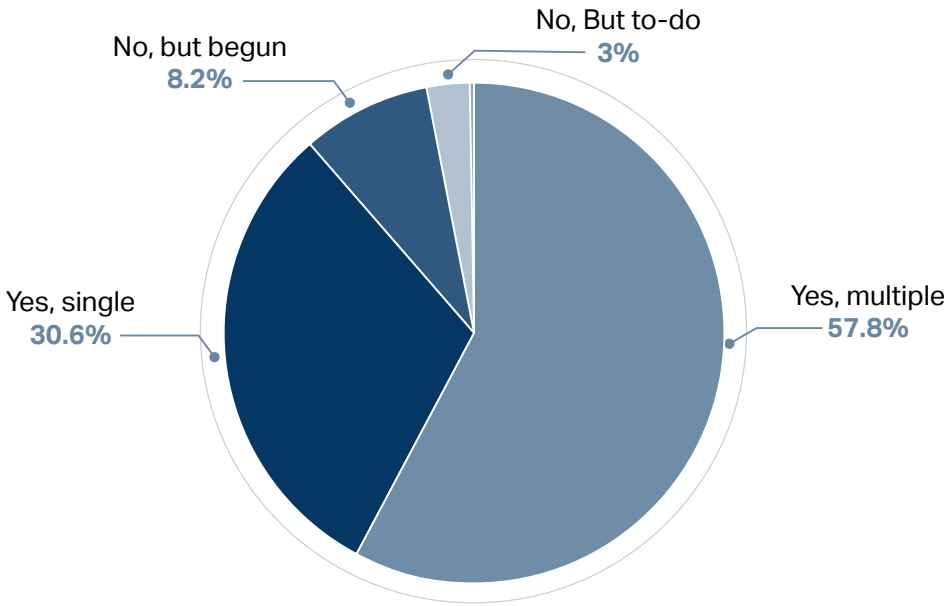
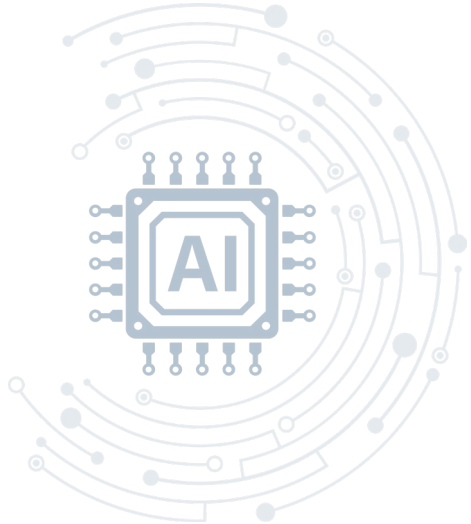


Figure 5 - Identifying Sources of Knowledge

[Q25] "Have you identified and assessed the existing pools of knowledge you'll use to power the AI?" (n=500)

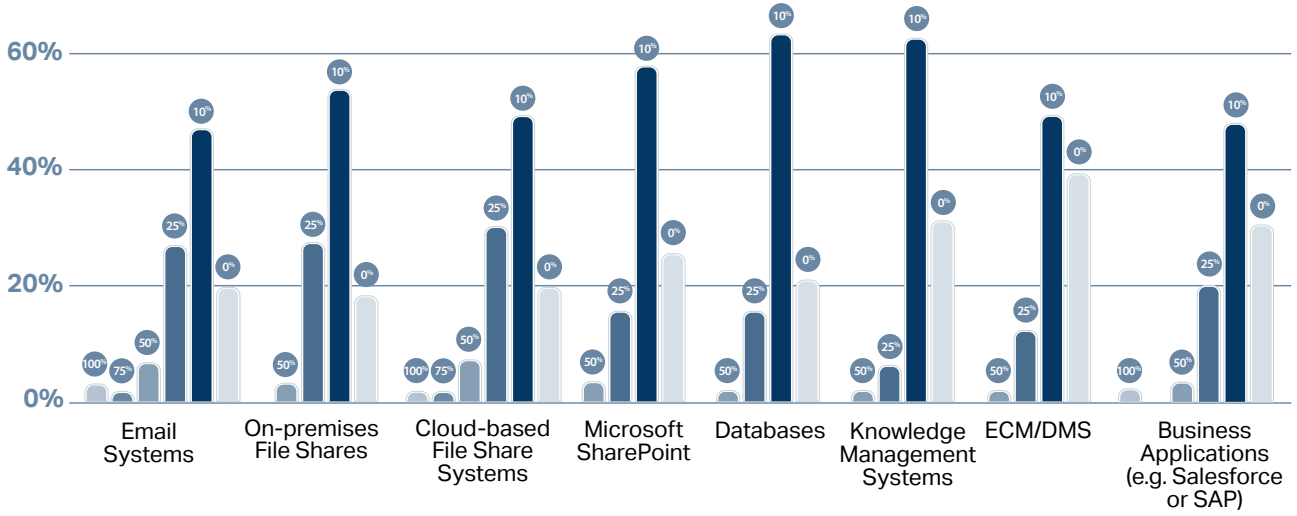


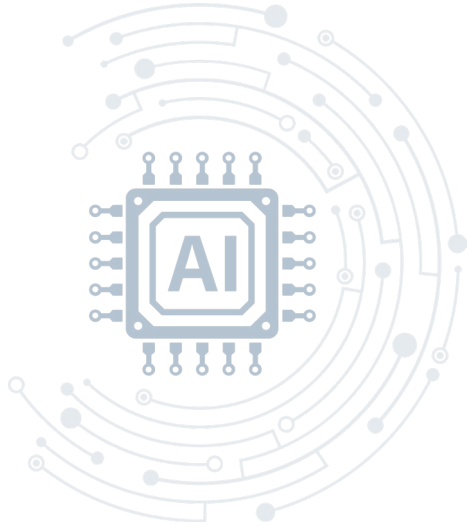


Adding perspective on the number of likely knowledge sources, respondents were also asked "At present, in which systems is your unstructured data stored/managed?" *(see Figure 6)*. The results revealed quite a disparate range of sources, each holding a relatively consistent volume.

Figure 6 - What Systems Store and Manage Unstructured Data

[Q4] "At present, in which system is your unstructured data stored/managed? Add an approximate percentage for each" n=(500)





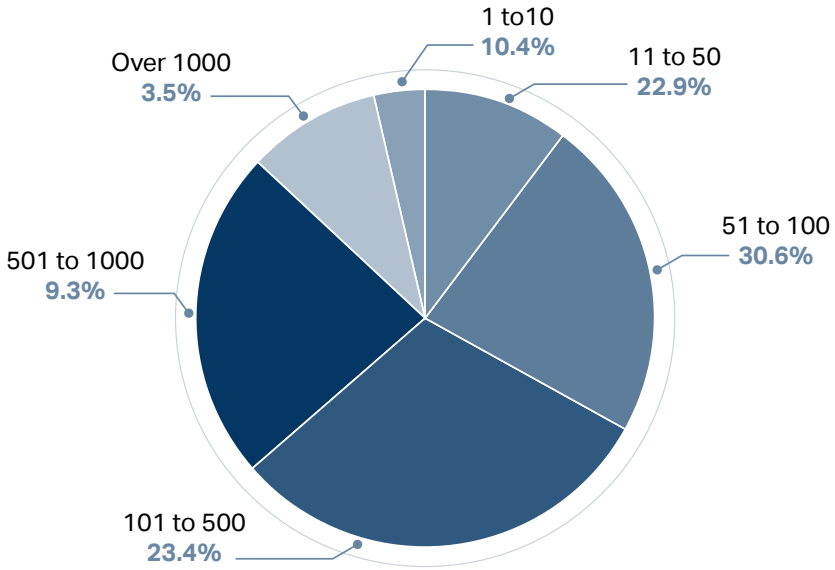
The strongest response across the board was that 10% of the total volume of unstructured data was managed by each of the suggested applications, with 25% broadly the second-highest response (along with 9% for systems the respondent said they didn't use at all). This set of responses suggests a less-than-scientific precision by the respondents; they know that they have some or all of these within their application estate and that they contain unstructured data of some sort. The result is a smoothing of the numbers, representing a somewhat hedged response.

In addition – referencing Q4 and Q25 in combination – respondents are aware of multiple systems containing unstructured data within their organizations, but only the 58% who responded that they have located multiple sources have potentially connected the dots.

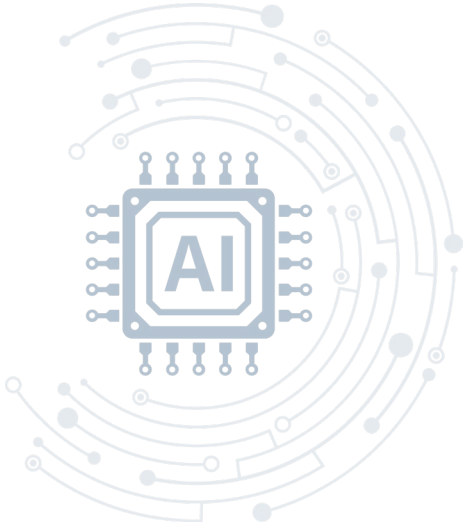
With regard specifically to Microsoft SharePoint, 75% of respondents said they used it within their organization [Q4]. Within that subset, 64% said they had up to 100 sites/instances, 36% said over 100, and 3% of those believed they had over 1,000 instances *(see Figure 7)*.

Figure 7 - Use of Microsoft SharePoint

[Q4a] "How many SharePoint sites/instances etc. do you have?" (n=376)



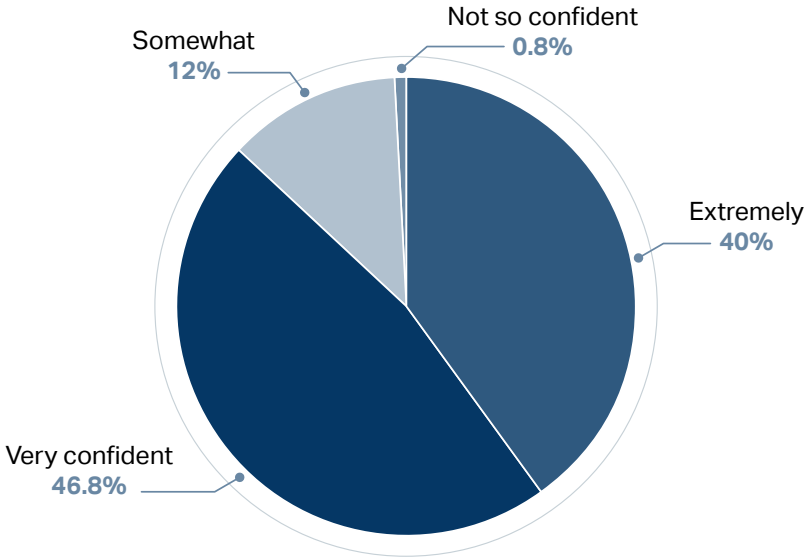
Looking at unstructured data within all sources, 99% of respondents said they were extremely, very, or somewhat confident in its quality and accuracy *(see Figure 8)*. However, only 40% are extremely confident.

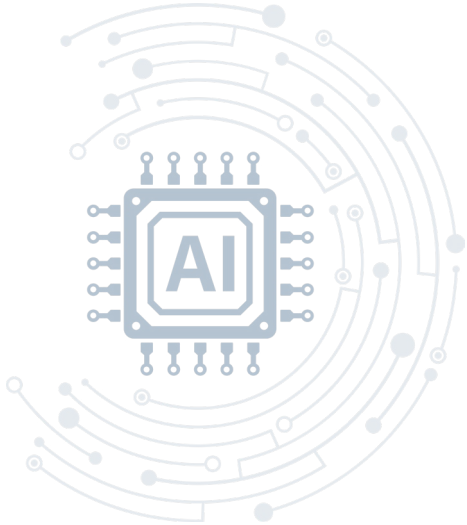


> *Microsoft 365 is a collection of tools optimized for creating and collaborating on content, yet organizations continue to try and use them for complex business processes. Moreover, using SharePoint to manage business critical data often leads to information chaos (duplicates, inability to find information), lack of process control and increased compliance risk that can be prevented with the right enterprise solution.*

Figure 8 - Confidence Level in Quality and Accuracy of Unstructured Data

[Q27] "How confident are you of the quality and accuracy of your internal sources of knowledge?" (n=500)

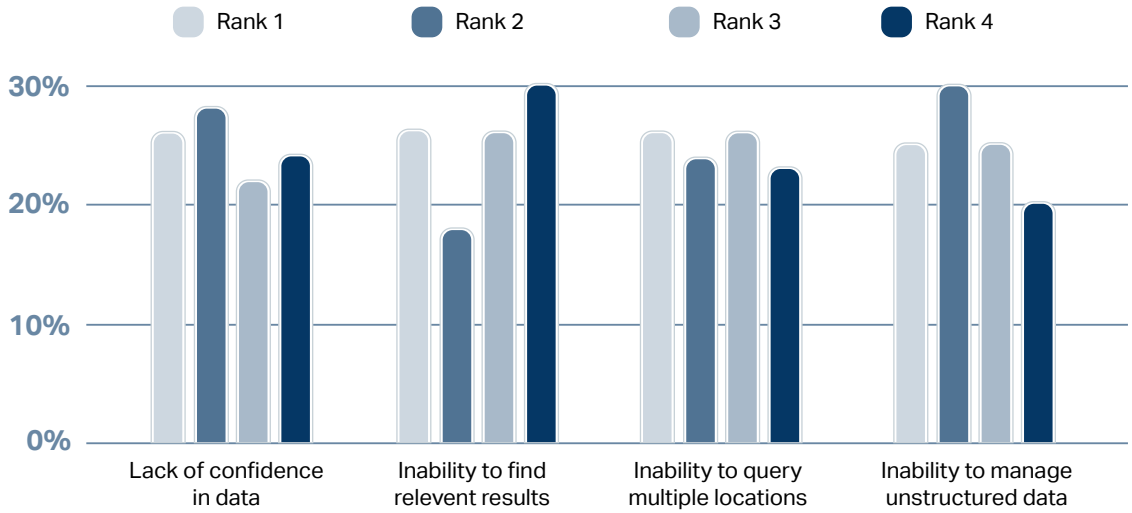


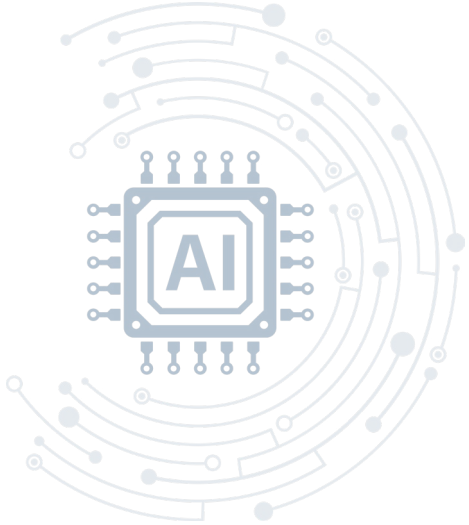


Respondents were also asked to rank suggested challenges around unstructured data on quality, findability, and management *(see Figure 9)*. Responses were smoothly blended across the board, without any significant spikes of extreme dissatisfaction (or any large pools of low dissatisfaction, which would suggest that, for example, findability has magically been solved).

Figure 9 - Challenges Related to Unstructured Data

"Please rank the challenges you face relative to unstructured data?" (n=485-490)





Cloud, costs, and even more software

Respondents estimated that about two-thirds of unstructured data is stored in the cloud vs. on-premises (67% median, 66% mean, standard deviation 19.84 [Q8]) albeit with a reasonable variance in responses. They also said that over half of overall storage is used for unstructured data (57% median, 56% mean, standard deviation 23.88 [Q9]). From these responses we can confirm that organizations have a lot of unstructured data, and around one-third of it is estimated to reside in-house.

It's worth remembering that unstructured data is far less likely than structured data to be closely and effectively managed directly within business applications, so the general march toward migrating those applications from on-premises to cloud does not reduce the size of the local unstructured volumes to the same extent as it does the structured.

When asked "Is your organization increasing, decreasing, or maintaining your current spend on managing unstructured data?" [Q11], 55% of respondents said that their costs were increasing (42% maintaining, 4% reducing). Those who said costs were increasing gave the following areas where the increased spend is going: 50% said software, 31% storage costs, 12% employees, and 7% external consulting *(see Figure 10)*.

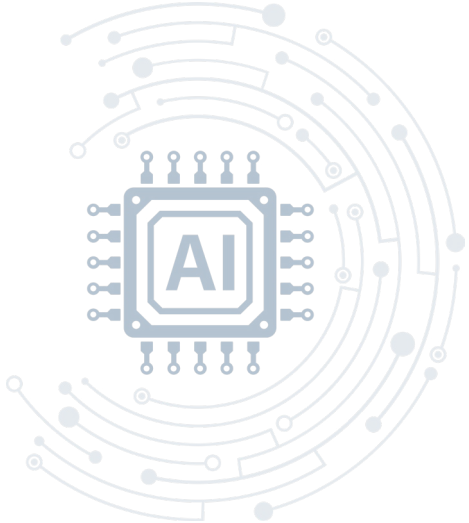
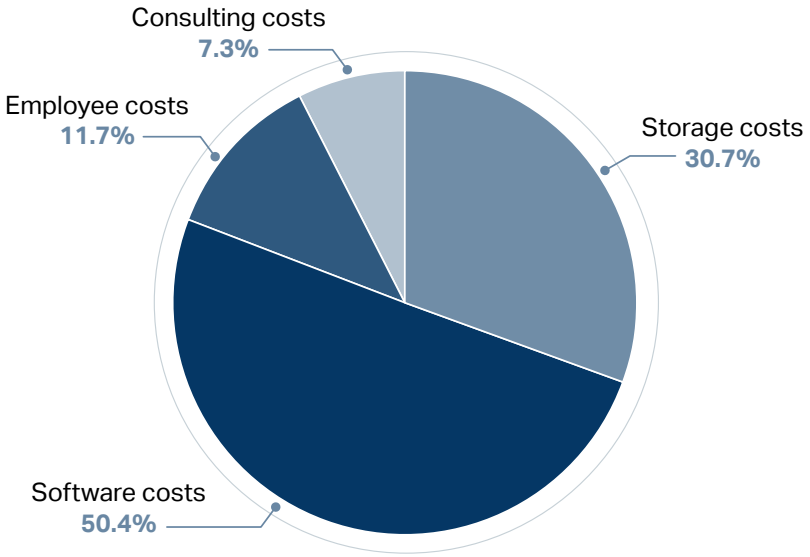


Figure 10 - Where Increased Cost of Managing Unstructured Data Is Going

[Q12] "Where is your organization primarily directing the increase in spending on managing unstructured data?" (n=274)



Respondents who said their spending was increasing were asked which software categories would receive the increased investment *(see Figure 11)*. Enterprise automation (69%) and knowledge management (61%) were the top two categories, ahead of AI (51%). It seems that unstructured data is useful for building AI applications, even if AI itself is less useful for managing that unstructured data.

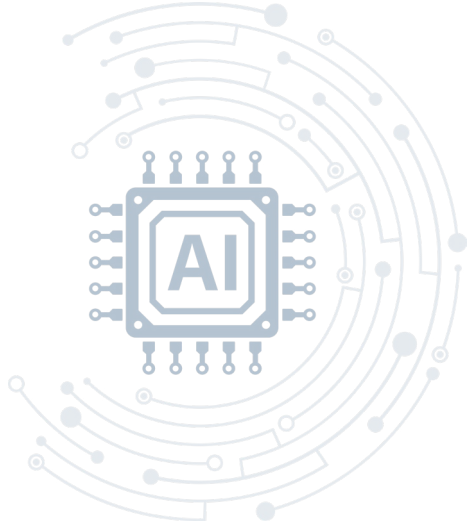
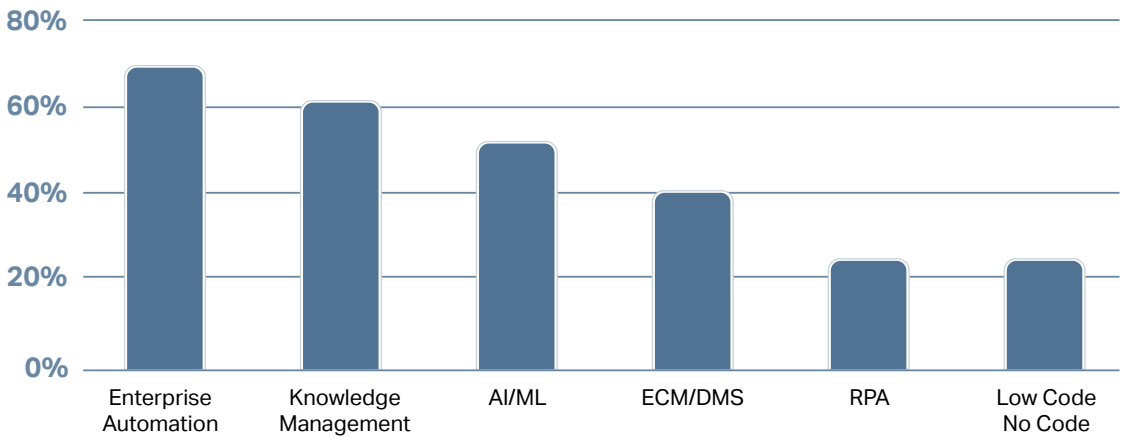


Figure 11 - Software Categories Receiving Increased Investment

[Q13] "Which software categories is your organization most focusing upon in terms of this increased investment?" (n=274)



Trends and Themes

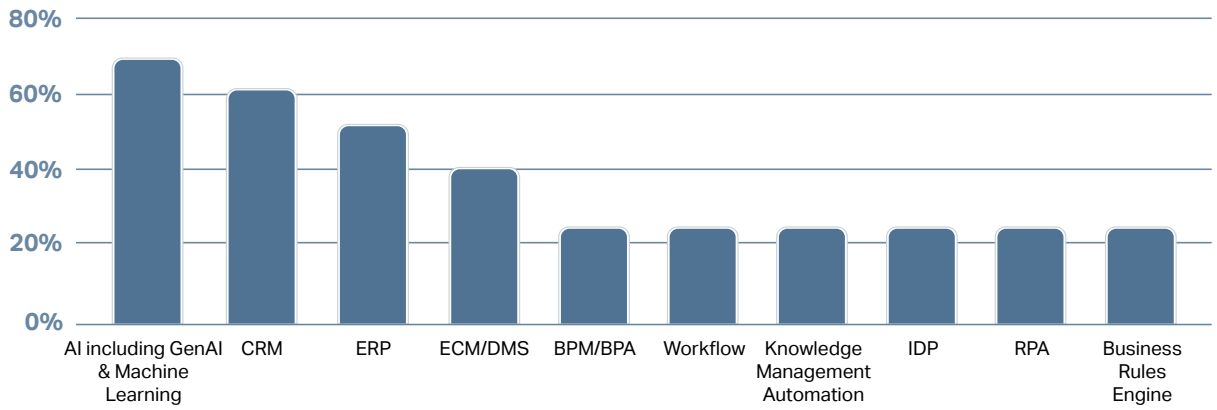
In addition to the prior analysis, trends and themes emerged in the survey data that are highlighted in this section.

Technology landscape

While it's not directly related to the specifics of unstructured data and AI, the survey also asked respondents which technologies they employ for process/work tasks within their organization *(see Figure 12)*. For a self-selecting sample responding to an invitation to an AI survey, that AI features so highly (76%, leading response) is unsurprising and also more or less validates the prior question about defining the AI state of the respondent's organization [S7]. IDP being referenced by 39% of respondents is perhaps higher than expected, and RPA with 32% perhaps a shade lower than expected.

Figure 12 - Technologies Employed for Process and Work Tasks

[Q3] "What technologies are currently employed to automate business processes and work-related tasks?" (n=500)



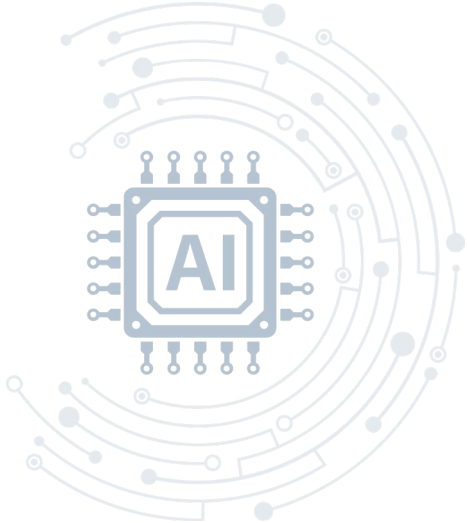
Paper vs. automation

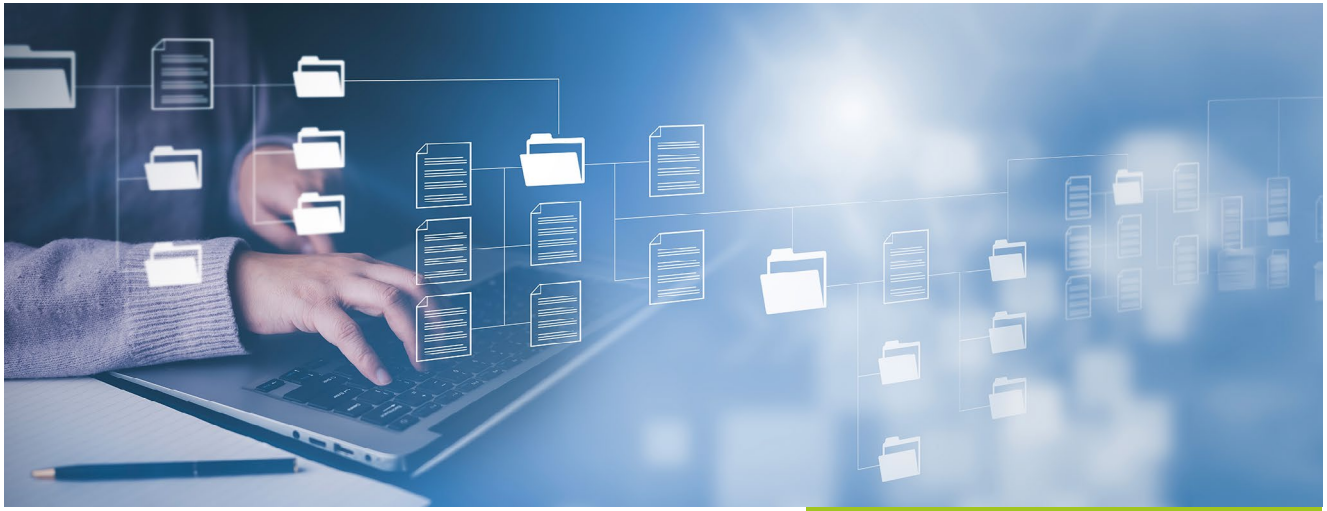
The survey contained two adjacent but unconnected questions on paper-based processes and fully automated processes.

One [Q15] asked, "What percentage of your organization's overall number of processes are currently paper-based?" Responses were 45% median, 46% mean, with 28.77 standard deviation.

The next [Q16] asked, "What percentage of your organization's processes are currently entirely automated?" Responses were 66% median, 63% mean, with 21.99 standard deviation.

In both cases, the breadth of responses was pretty wide (from 0-100%) and given the breadth of the question, chances are no single response was accurate. However, it is an interesting contrast to see how much paper there still is across business functions and how full automation – while more common – is still far less prevalent than many had predicted a decade ago would now be the case. The paper number in particular suggests a potential reason for the higher-than-expected IDP response referenced in the previous trend.





Line-of-business focus

An earlier section of this paper, "Processes and tasks – indicators of AI readiness," discussed the importance of task and process identification for AI readiness. Survey respondents were asked additional questions that provide interesting insights.

Asked whether the processes were internal- or external-facing [Q20], 24% said internal-only; 12% external-only; and 61% said a mix.

Participants were also asked about the department-level focus of that work, and possible responses were "current," "planned," "likely focus," and "not applicable/no focus" *(see Figure 13)*.

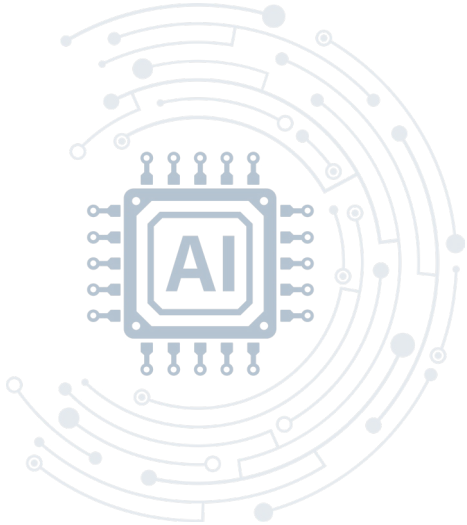
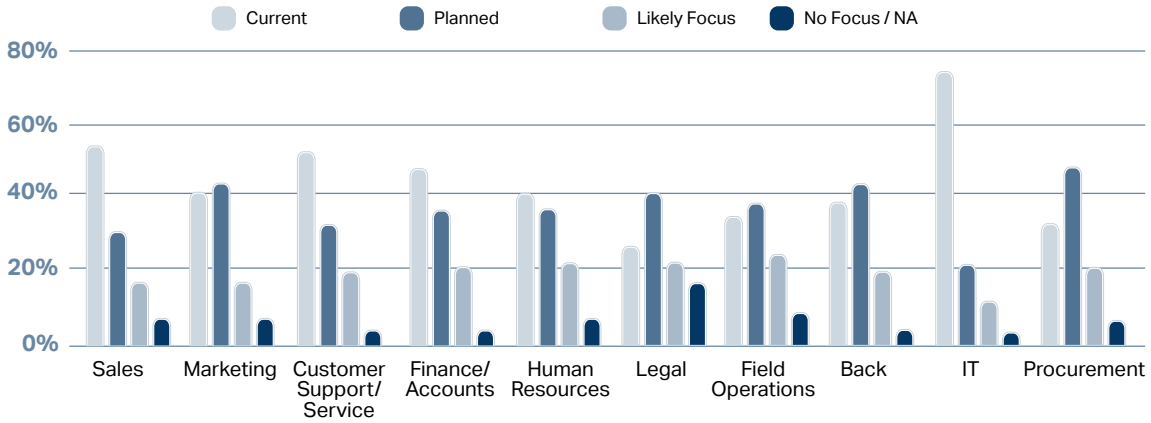


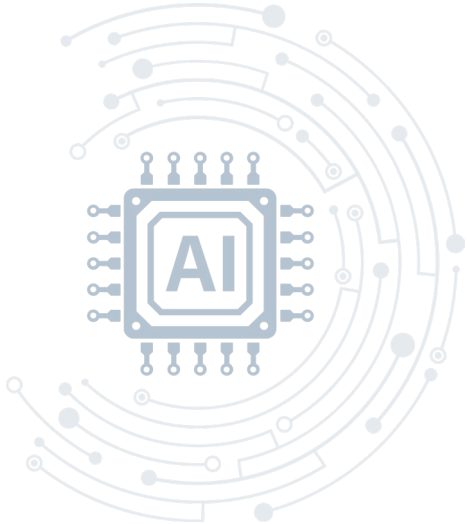
Figure 13 - Task and Process Identification for AI Readiness: Lines of Business

[Q21] "In identifying business processes for improvement are you focused on any of the following departments?" (n=500)



Of lines-of-business with "current" focus, IT (68%), Sales (50%) and Customer Support (49%) were the strongest responses. This is reflective of the fact that code/development support use cases are among the most developed for GenAI/LLM offerings, and that support for Sales and Customer Support have been the most common copilot assistant use cases from the large-productivity-suite software vendors to date. (Pre-survey, we expected Marketing to be among these, but the data suggests that it's "in the pack" right now).

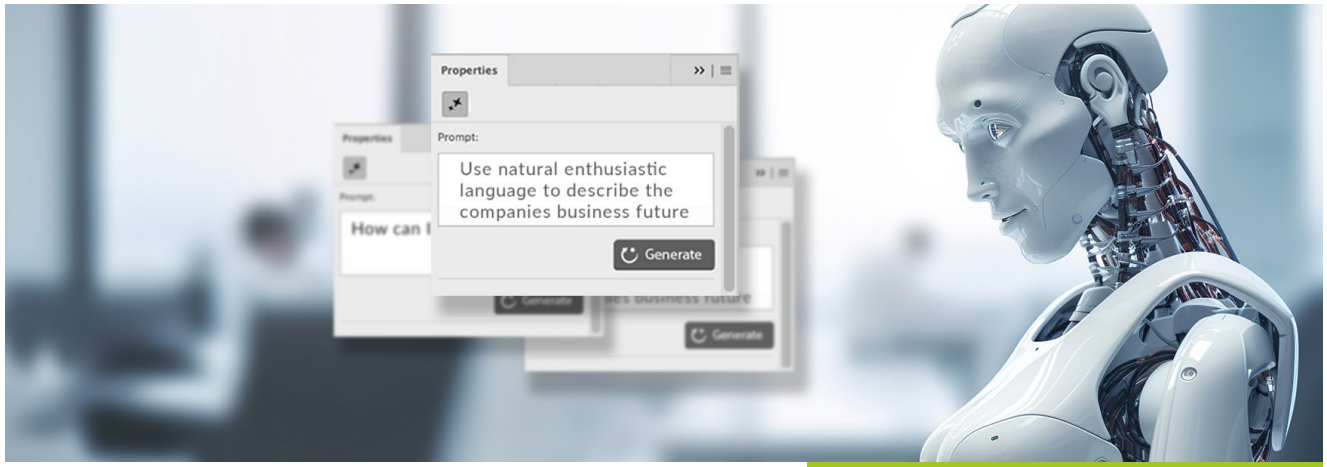
Among "planned" or "likely" focus, results are a smooth blend suggesting not much has been ruled out, but there's no overwhelming case for any to be fast-tracked for investment. Within not applicable/no focus, there is something of an overscore for Legal (16%, against a mean of 6%) which is worth noting when set against some suggestions that this area is ripe for this type of technological transformation.



Horizontal vs. vertical AI

As briefly referenced in the "Operational, accelerated, transformational: current use of AI" section earlier, the trends around AI adoption are closely aligned across the five verticals included in this study. That is not to say that current AI technology is suitable for all industries, but rather the current state of play – especially regarding the explosion of GenAI-derived product offerings in the last 24 months – has been directed toward horizontal functionality. In practice, this means that organizations are being offered desktop functionality for common productivity tasks (for example, copilot assistants) and where specialisms appear, they are directed toward the suggested needs of common line-of-business operations (sales, marketing, customer services).

This is why there is no great leap as yet from one vertical or another toward specific adoption, and perhaps there won't be until smaller, specialist language – and potentially, action – models make a bigger impression on the marketplace. This will most likely happen via copilots (probably through OEM deals for vendors specializing in industry models for the likes of healthcare, pharmaceuticals, and similar industries).



Topics for Further Study

A successful research study tends to elicit follow-up questions that were either not included in or out of scope for the original work, and ***Market Momentum Index: AI and Unstructured Data Management*** is no different. While it has uncovered the degree to which organizations are developing and employing AI applications in significant volumes, concerned about security and accuracy, and investing in understanding tasks and processes, there is still more that would be beneficial to understand. For example:

- Breaking down projects in production by vertical, line-of-business
- Understanding the timeline from conception to launch
- Surveying the use of in-house development, open source, and commercial AI projects/ platforms/products
- Benchmarking the use of RAG and other unstructured data augmented tools for AI
- Understanding the budgetary source(s) for AI projects both in production and planned
- Breaking down the roles for knowledge management within the development and ongoing support of those RAG-derived AI projects in production

It is hoped that a future, updated version of ***Market Momentum Index: AI and Unstructured Data Management*** will be able to include some or all of these areas.

Conclusions and Recommendations

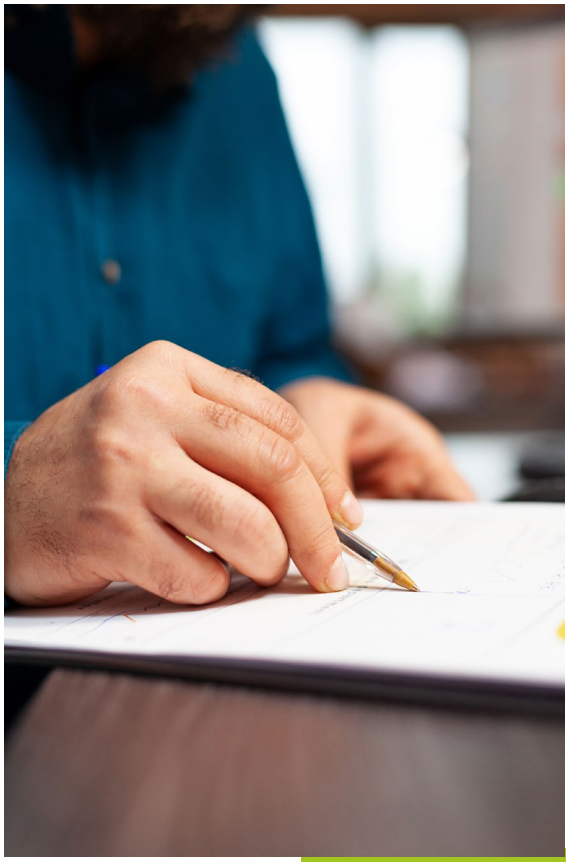
This first version of *Market Momentum Index: AI and Unstructured Data Management*, based on research conducted in mid-2024, provides strong indications of significant momentum for organizations to develop and launch AI applications: 77% of respondents have AI projects either in production or currently in evaluation, with 79% of those in production.

- It is important for the software industry that has invested so much in betting on AI (especially, of late, generative AI) to continue this momentum, supporting those already on the train and bringing more on board in the immediate future. To this end, Deep Analysis has recommendations for both customers and the industry that supports them.

- It's heartening that so many organizations are employing formal process analysis practices – including extensive use of process mining – into their AI projects. For AI to pay its way in enterprises, it has to be woven into the fabric of how those organizations work; enhancing, augmenting, and ultimately enabling ways of working that previously were impossible or impractical. Process mining, along with its close relation task mining, when used in conjunction with teams of practitioners provides a strong insurance policy against developing and deploying AI applications that do not provide beneficial fit and function to organizations. Deep Analysis recommends that AI-advocating software vendors put process analysis – including the use of process and task mining technology – more centrally in their go-to-market.

- The primary AI concerns among respondents were directed toward performance/accuracy and security/compliance. Both of those subjects have increasingly featured in the sales pitches of the industry, even though at times that's been somewhat in the wash of the trend toward using the spurious term "hallucinations" (which for all sorts of reason is entirely inappropriate). Certainly, both topics become more complex and pressing when considered in a whole that includes not only LLMs, but also organizational-specific data which may be destined to a RAG solution or similar. Ensuring that this data is assured for quality prior to its inclusion within AI applications is going to be a partnership between software vendors, systems integrators, and customers alike. This is true especially for unstructured data, which is not only essential to RAG, etc., but tends to be less well-understood, classified, and adequately secured. Deep Analysis recommends that as part of recognizing enterprise concerns about performance/accuracy and security/compliance, practicality should override hyperbole in how the industry speaks to these organizations.

- What's coming down the track is much larger-scale data integration as those RAG solutions become more standardized and essential in delivering the sort of organizational-specific AI projects that are necessary to maintain momentum over time. There's an indication within the results of this survey that while organizations have some idea of where unstructured data might be lurking and might have even secured and audited some of it, the extent of those volumes is probably less well understood. The longer that is the case, the more inhibited organizations will be in their ability to make use of that data (both natively electronic and from paper-based processes) to develop future AI-derived use cases. Nothing is likely to choke the momentum of AI faster than a lack of trust in or a shortage of appropriate proprietary data, and it's in everybody's interest to ensure that its free, fresh flow continues. Deep Analysis recommends that organizations seeking to begin their exploration or continue their journey into the operational use of AI get a broader handle on the extent of their existing unstructured data – both its volume and its veracity. Failing to do so will quickly inhibit their ability to gain the most from this potentially deep technological change.



Appendix:

Survey Answer Tables

[S2] Which of the following best describes your organization's industry?	Total (Number)	Total (Percentage)
Total	500	100%
Financial and Insurance	100	20%
Professional, Scientific and Technical Services	100	20%
Manufacturing Sector	100	20%
Mining, Quarrying, and Oil & Gas Extraction	100	20%
Utilities Sector	100	20%
Other	0	0%

[S3] What was the annual revenue for your organization in fiscal 2023?	Total (Number)	Total (Percentage)
Total	500	100%
Less than $100K	0	0%
$100K - $999K	0	0%
$1M - $9.9M	44	9%
$10M - $49.9M	75	15%
$50M - $99.9M	90	18%
$100M - $499.9M	108	22%
$500M - $999.9M	108	22%
$1B+	75	15%
Don't know	0	0%

Survey Answer Tables

[S4] Please provide your best estimate of the total employees at your organization worldwide. Please include all subsidiaries, divisions, and branches worldwide.	Total (Number)	Total (Percentage)
Total	500	100%
1 - 99 Employees	0	0%
100 - 499 Employees	0	0%
500 - 999 Employees	164	33%
1,000 - 1,499 Employees	122	24%
1,500 - 1,999 Employees	63	13%
2,000+ Employees	151	30%
Don't know	0	0%

[S5] What level best represents your role within your organization?	Total (Number)	Total (Percentage)
Total	500	100%
Executive/C-level (CEO, CIO, CTO, CDO, etc.)	184	37%
VP-level	26	5%
Director-level	150	30%
Manager-level	140	28%
Supervisor or team-level	0	0%
Individual contributor	0	0%
Other	0	0%

Survey Answer Tables

[S6] Which of the following most closely represents the department/function you work in?	Total (Number)	Total (Percentage)
Total	316	100%
IT Organization	165	52%
Digital Transformation/Innovation Office	16	5%
R&D	3	1%
Product Development/Engineering	18	6%
Manufacturing/Operations	92	29%
Corporate Strategy & Planning	22	7%
Facilities Management	0	0%
Marketing/Sales	0	0%
Human Resources	0	0%
Finance	0	0%
Other	0	0%

Survey Answer Tables

[S7] As it relates to the use of AI (including GenAI and/or LLMs) within the business, at what stage would you place your organization?	Total (Number)	Total (Percentage)
Total	500	100%
Concerned - Risk adverse, searching for where this can go wrong	25	5%
Skeptical - Wait and see	20	4%
Motivated - Excited about the technology but haven't used it yet, seeking use-cases and planning to pilot and then deploy where possible	68	14%
Experimenting - Experimenting but haven't deployed any large-scale projects yet	78	16%
Operational - We have begun to adopt AI into some of our day-to-day functions	107	21%
Accelerated - Adopting AI into many of our day-to-day functions, continuously seeking new use cases, and deploying where possible	130	26%
Transformational - AI is built into the DNA of our business and into almost all our day-to-day functions. We rely on AI to do some heavy lifting for the business and as a value generator for our customers	72	14%
Don't know		0

Survey Answer Tables

[Q3] What technologies are currently employed to automate business processes and work-related tasks?	Total (Number)	Total (Percentage)
Total	500	100%
BPM/BPA - Business Process Management/Automation	211	42%
RPA - Robotic Process Automation	158	32%
ECM - Enterprise Content Management/DMS Document Management System	212	42%
Knowledge Work Automation	198	40%
CRM - Customer Relationship Management	324	65%
ERP - Enterprise Resource Planning	241	48%
Business Rules Engine	123	25%
Workflow Orchestration	199	40%
IDP - Intelligent Document Processing	197	39%
AI - Artificial Intelligence including GenAI & Machine Learning	379	76%
Other	3	1%

Survey Answer Tables

[Q4] At present, in which systems is your unstructured data stored/managed? Add an approximate percentage for each																
	Email Systems		On-Premises File Shares		Cloud-based File Share Systems		Microsoft SharePoint		Databases		Knowledge Management Systems		ECM/DMS		Business Applications (e.g. Salesforce or SAP)	
Total	500	100%	500	100%	500	100%	500	100%	500	100%	500	100%	500	100%	500	100%
100%	8	2%	0	0%	5	1%	1	0%	1	0%	0	0%	0	0%	3	1%
75%	3	1%	2	0%	4	1%	0	0%	0	0%	0	0%	0	0%	0	0%
50%	30	6%	9	2%	36	7%	13	3%	5	1%	4	1%	3	1%	9	2%
25%	130	26%	135	27%	149	30%	74	15%	77	15%	29	6%	58	12%	100	20%
10%	231	46%	266	53%	246	49%	288	58%	313	63%	310	62%	245	49%	238	48%
0%	98	20%	88	18%	60	12%	124	25%	104	21%	157	31%	194	39%	150	30%

[Q4a] How many SharePoint sites/instances etc. do you have?	Total (Number)	Total (Percentage)
Total	376	100%
1 to 10	39	10%
11 to 50	86	23%
51 to 100	115	31%
101 to 500	88	23%
501 to 1000	35	9%
Over 1000	13	3%

Survey Answer Tables

[Q5] Do you currently operate a data retention policy?	Total (Number)	Total (Percentage)
Total	500	100%
Yes and it is strictly adhered to	376	75%
Yes but it is not strictly adhered to	107	21%
Not sure	7	1%
No	7	1%
Unfamiliar with the concept of a "retention policy"	3	1%

[Q6] Is this retention policy to meet an industry compliance standard?	Total (Number)	Total (Percentage)
Total	500	100%
Yes	401	80%
No	19	4%
Partly (compliance is just one factor)	69	14%
Unsure	11	2%

[Q7] Are specific costs recorded for the operation of this retention policy?	Total (Number)	Total (Percentage)
Total	500	100%
Yes	420	84%
No	52	10%
Unsure	28	6%

Survey Answer Tables

[Q8] What proportion of your organization's unstructured data resides in the cloud versus on-premises?	Total (Number)	Total (Percentage)
Total	500	100%
Mean	65.52	
Median	67.00	
Standard Deviation	19.84	

[Q9] What percentage of your organization's overall storage usage is for unstructured data?	Total (Number)	Total (Percentage)
Total	500	100%
Mean	56.46	
Median	56.50	
Standard Deviation	23.88	

[Q10] Is your organization increasing, decreasing, or maintaining its current storage volumes?	Total (Number)	Total (Percentage)
Total	500	100%
Decreasing	19	4%
Maintaining	186	37%
Increasing	295	59%

Survey Answer Tables

[Q11] Is your organization increasing, decreasing, or maintaining your current spend on managing unstructured data?	Total (Number)	Total (Percentage)
Total	500	100%
Decreasing	18	4%
Maintaining	208	42%
Increasing	274	55%

[Q12] Where is your organization primarily directing the increase in spending on managing unstructured data?	Total (Number)	Total (Percentage)
Total	274	100%
Storage Costs	84	31%
Software	138	50%
Employee Costs	32	12%
External Consulting Costs	20	7%

Survey Answer Tables

[Q13] Which software categories is your organization most focusing upon in terms of this increased investment?	Total (Number)	Total (Percentage)
Total	274	100%
ECM/DMS	109	40%
Low Code/No Code Applications	65	24%
Enterprise Automation (for example, BPA & BPM)	190	69%
Knowledge Management	166	61%
RPA	66	24%
AI/ML	141	51%
Other	1	0%

[Q14] Please rank the challenges you face relative to unstructured data?								
	Lack of confidence in the accuracy of unstructured data (version control, out of date)		Inability to find relevant results		Inability to query across multiple storage locations		Inability to effectively manage unstructured data volumes	
Total	485	100%	489	100%	490	100%	490	100%
1	124	26%	127	26%	128	26%	121	25%
2	136	28%	88	18%	119	24%	147	30%
3	109	22%	126	26%	128	26%	124	25%
4	116	24%	148	30%	115	23%	98	20%

Survey Answer Tables

[Q15] What percentage of your organization's overall number of processes are currently paper-based?	Total (Number)	Total (Percentage)
Total	500	100%
Mean	46.27	
Median	45.00	
Standard Deviation	28.77	

[Q16] What percentage of your organization's processes are currently entirely automated?	Total (Number)	Total (Percentage)
Total	490	100%
Mean	62.98	
Median	66.00	
Standard Deviation	21.99	

[Q17] If your organization uses, or plans to use, AI (including GenAI) do you believe this will impact how you currently manage unstructured data?	Total (Number)	Total (Percentage)
Total	500	100%
Yes	448	90%
No	20	4%
Unsure	32	6%

Survey Answer Tables

[Q18] Have you identified processes in your organization that you think AI will help improve?	Total (Number)	Total (Percentage)
Total	500	100%
Yes, we have identified several processes	317	63%
Yes, we have identified at least one process	142	23%
No, but we have begun attempting to identify processes	26	5%
No, but it is on our to-do list	7	1%
No and we do not think it is necessary at this stage	8	2%

[Q19] Do you plan to leverage AI for any or both of the following?	Total (Number)	Total (Percentage)
Total	500	100%
Generative AI - produce something from scratch	109	22%
Extractive AI - produce something based on existing content	82	16%
Both generative and extractive AI	309	62%

[Q20] In identifying business processes for improvement are you focused on back or front office operations?	Total (Number)	Total (Percentage)
Total	500	100%
Yes, internal (back office) facing processes	122	24%
Yes, external (front office) facing processes	60	12%
Yes, a mix of both internal and external facing processes	306	61%
No specific focus	12	2%

Survey Answer Tables

[Q21] In identifying business processes for improvement are you focused on any of the following departments?																				
	Sales		Marketing		Customer Support / Service		Finance / Accounts		Human Resources		Legal		Field Operations		Back Office Operations		IT		Procurement	
Total	500	100%	500	100%	500	100%	500	100%	500	100%	500	100%	500	100%	500	100%	500	100%	500	100%
Current	251	50%	188	38%	244	49%	219	44%	192	38%	127	25%	165	33%	184	37%	341	68%	151	30%
Planned	141	28%	205	41%	151	30%	169	34%	168	34%	190	38%	180	36%	206	41%	101	20%	221	44%
Likely focus	76	15%	78	16%	91	18%	99	20%	104	21%	104	21%	117	23%	92	18%	48	10%	99	20%
Not applicable/ No focus	32	6%	29	6%	14	3%	13	3%	36	7%	79	16%	38	8%	18	4%	10	2%	29	6%

[Q22] Are you planning to use any specific process identification software tools (e.g. Process Mining) to assist you?	Total (Number)	Total (Percentage)
Total	500	100%
Yes, we're using Process Mining	220	44%
Yes, we're using process mapping tools other than Process Mining	203	41%
No, we've used manual process mapping, interviews, and whiteboarding	38	8%
No, we've not chosen to use any specialist process identification software, but we know they exist	26	5%
We are unaware of Process Mining technologies	13	3%

Survey Answer Tables

[Q23] Which other software tools will you be using as part of your AI project work?	Total (Number)	Total (Percentage)
Total	500	100%
RPA suites	108	22%
Productivity tools (e.g. spreadsheets, project planning)	318	64%
BPM/BPA systems	170	34%
Document/Records/Knowledge Management platforms	265	53%
Business Applications (ERP/CRM, etc.)	294	59%
Data Management	340	68%
Business Intelligence tools	256	51%

[Q24] Do you know the specific tasks within those processes that AI will power?	Total (Number)	Total (Percentage)
Total	500	100%
Yes, we have identified a number of tasks	313	63%
Yes, we have identified at least one task	121	24%
No, but we have begun attempting to identify tasks	49	10%
No, but it is on our to-do list	13	3%
No and we do not think it is necessary at this stage	4	1%

Survey Answer Tables

[Q25] Have you identified and assessed the existing pools of knowledge you'll use to power the AI?	Total (Number)	Total (Percentage)
Total	500	100%
Yes, we have identified a number of knowledge bases	289	58%
Yes, we have identified at least one knowledge base	153	31%
No, but we have begun attempting to identify knowledge bases	41	8%
No, but it is on our to-do list	15	3%
No and we do not think it is necessary at this stage	2	0%

[Q26] How important do you think existing internal sources of knowledge will be to the success of your AI?	Total (Number)	Total (Percentage)
Total	500	100%
Mean	77.03	
Median	80.00	
Standard Deviation	17.13	

[Q27] How confident are you of the quality and accuracy of your internal sources of knowledge?	Total (Number)	Total (Percentage)
Total	500	100%
Extremely confident	200	40%
Very confident	234	47%
Somewhat confident	60	12%
Not so confident	4	1%
Not at all confident	2	0%

Survey Answer Tables

[Q28] What prevents your enterprise from effectively leveraging AI?	Total (Number)	Total (Percentage)
Total	500	100%
AI Hallucinations - Concern about AI systems generating unrelated, incorrect, or misleading information	157	31%
Security and Access Compliance - Challenges in ensuring AI systems meet strict security and access control requirements	213	43%
Data Integration and Silos - Difficulties in curating and integrating data from disparate storage systems and silos	194	39%
Cost of Implementation - High financial investment required for AI deployment, including infrastructure, maintenance, and talent acquisition	194	39%
Scalability Issues - Concerns about the ability of AI solutions to scale with the growth of the enterprise	141	28%
Lack of Skilled Personnel - Shortage of qualified professionals to develop, implement, and manage AI technologies	167	33%
Ethical and Bias Concerns - Worries about ethical implications and the potential for AI systems to exhibit biased behavior	126	25%
Regulatory Compliance - Difficulties in ensuring AI systems comply with industry-specific regulations and standards	155	31%
User and Stakeholder Adoption - Resistance from employees, stakeholders, or customers to adopting AI technologies	109	22%
Integration with Existing Systems - Challenges in integrating AI solutions with current IT infrastructure and legacy systems	191	38%
Performance and Accuracy - Concerns about the reliability, performance, and accuracy of AI models in critical applications	202	40%
Transparency and "Explainability" - The need for AI systems to be transparent and their decision-making processes to be easily explainable	163	33%

Survey Answer Tables

[Q29] Where additional skills are required, will you acquire them by training existing staff, hiring, or combining both?	Total (Number)	Total (Percentage)
Total	500	100%
We will look to externally hire the required skills	201	40%
We're planning to train existing staff to take on the skills that we currently lack	153	31%
We'll use a mixture of training and external hiring to fill in our missing skills	146	29%

[Q30] Where will you look for external help if required?	Total (Number)	Total (Percentage)
Total	500	100%
System Integrators (consultants) we already work with	157	31%
New System Integrators and Consultants	191	38%
Existing external technology vendor resources	100	20%
Ad-hoc independent consultancy as required	24	5%
We have sufficient internal resources to do the work	28	6%

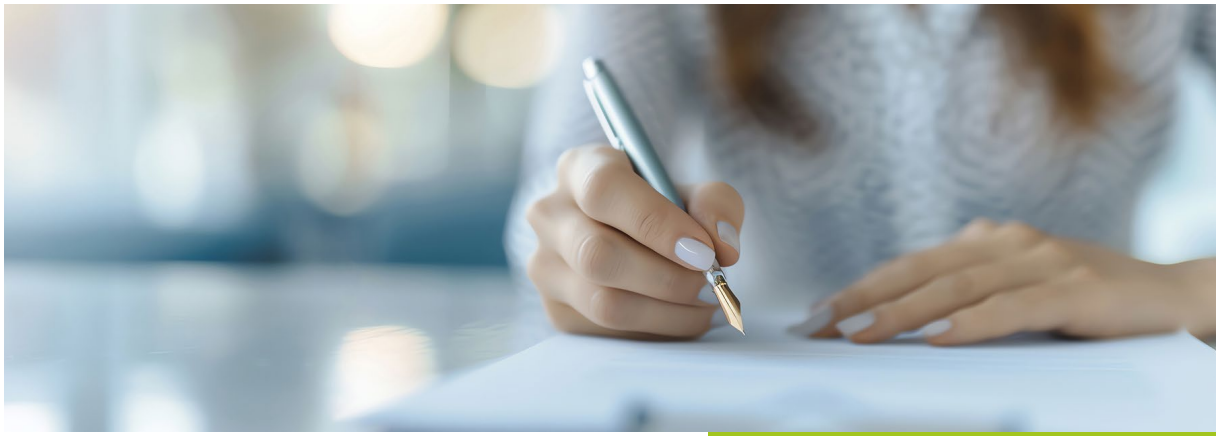
Survey Answer Tables

[Q31] Have you agreed on the measurements that you'll use to judge the success of your AI?	Total (Number)	Total (Percentage)
Total	500	100%
Yes, we have identified a number of key measurements that we will employ	265	53%
Yes, we have identified at least one key metric we will be using	167	33%
No, but we have begun attempting to identify which measurements we will use	50	10%
No, but it is on our to-do list	13	3%
No and we do not think it is necessary	5	1%

[Q32] Are these planned success measurements likely to be just for the implementation phase of the project or for the ongoing production AI system?	Total (Number)	Total (Percentage)
Total	500	100%
For project stage sign-offs	65	13%
We are planning for measurements to continue once the project is in production	278	56%
We are planning for different measurements to be used in production than are used in the implementation project	128	26%
We are only planning for measurements to be used once in production	15	3%
We have no firm plans yet in place	14	3%

Survey Answer Tables

[Q33] If you are embarking on an automation project – what are your primary business drivers (pick 1 to 3 in order of priority)?																										
	To reduce operating costs		To automate manual process activities		Reduce headcount		To increase business agility		Improve the quality of products/ services		To remain competitive		Grow revenue		Standardize operations		Meet customer experience demands		Simplify operational complexity		Meet regulatory compliance requirements		To increase the speed of current business process activities		Expand our business into new areas	
Total	151	100%	128	100%	41	100%	138	100%	170	100%	117	100%	166	100%	76	100%	106	100%	118	100%	82	100%	119	100%	88	100%
1	53	35%	34	27%	16	39%	43	31%	58	34%	35	30%	78	47%	23	30%	26	25%	38	32%	28	34%	41	34%	27	31%
2	45	30%	46	36%	10	24%	50	36%	61	36%	45	38%	44	27%	31	41%	32	30%	44	37%	20	24%	41	34%	31	35%
3	53	35%	48	38%	15	37%	45	33%	51	30%	37	32%	44	27%	22	29%	48	45%	36	31%	34	41%	37	31%	30	34%



About Deep Analysis

Deep Analysis

Deep Analysis is a leading industry analyst firm specializing in providing innovative insights and strategic advisory services for enterprises navigating the complexities of digital transformation. With a focus on emerging technologies, including AI, automation, blockchain, and data management, Deep Analysis delivers in-depth research, market analysis, and actionable recommendations tailored to the unique needs of businesses across various industries. Our team of seasoned analysts leverages a wealth of experience and a global perspective to empower organizations with the knowledge and foresight needed to make informed decisions, drive innovation, and maintain a competitive edge in an ever-evolving market landscape.

For further information visit:

www.deep-analysis.net

About AIIM

The Association for Intelligent Information Management (AIIM)

Founded in 1944, the Association for Intelligent Information Management (AIIM) is a nonprofit organization serving information leaders in over 67 countries worldwide. AIIM's vision is to create a world where every organization benefits from intelligent information and data management to achieve better business outcomes. AIIM helps information leaders manage and prepare unstructured data for AI and automation by providing advice, certification, training, and peer-to-peer support. Through practical and approachable resources, AIIM enables organizations to leverage their information assets effectively, ultimately leading to improved business performance and success.

For more information visit:

www.aiim.org

About M-Files

The M-Files knowledge work automation platform harnesses the power of generative artificial intelligence (GenAI) and automation to boost the performance of knowledge workers. M-Files automates the entire process from document creation and management to workflow automation, external collaboration, enterprise search, security, compliance and audit trail.

M-Files' GenAI assistant, M-Files Aino, enables real-time, context-aware assistance to process vast amounts of structured and unstructured information quickly, get answers to complex questions and arrive at conclusions faster. Its multilingual capabilities support global enterprises by summarizing documents and answering questions in any language, saving these insights as searchable metadata.

The M-Files metadata foundation drives superior AI experiences by creating a unique, customer-specific information model that surrounds all content, ensuring safe and high-quality results.

With M-Files, organizations automatically get the mandatory enablers for successful AI deployment: connectivity, confidentiality and curation.

- **Connectivity:** M-Files seamlessly integrates data from all key storage systems and software solutions and automatically classifies it, providing AI with access to the high-quality data inputs it needs to deliver real value.
- **Confidentiality:** M-Files prioritizes information security, ensuring that only authorized users access sensitive data. Its advanced metadata tagging enables granular access controls, meeting the needs of security-conscious clients.
- **Curation:** M-Files ensures AI uses relevant and current information with automated version control, content governance and precise business context. Placing content in context improves search capabilities and AIs ability to deliver more accurate results compared to traditional approaches.

To request a demo or download a free 30-day trial of M-Files visit:

www.m-files.com

aiim
Association for Intelligent Information Management

+1 301 587 8202
hello@aiim.org
www.aiim.org

Research by Deep Analysis in partnership with AIIM

Deep Analysis

aiim
Association for Intelligent Information Management

Underwritten by

M-Files