
# Organizational Readiness for Generative Artificial Intelligence

## Leveraging Unstructured Data for Success

Sponsored by

Hyland™

aiim
Association for Intelligent Information Management

AVAILTEK

# 1. Executive Summary

Most leaders recognize the impact that generative artificial intelligence (GenAI) can have on employee productivity and organizational performance. However, those benefits aren't guaranteed. Organizations that want to take advantage of GenAI must create the enabling conditions for success, or initiatives may be a waste of resources with little to no return on investment.

Unstructured data (content) is king, and it is at the core of enabling GenAI. This readiness whitepaper is designed to help organizations leverage their assets for AI success.

## 1.1. Understanding Content's Role in GenAI

As used here, "content" is defined as structured and unstructured data. Structured data, with rows, columns, and relationships, makes up the heart of database and operational systems. Unstructured data, such as emails, documents, and multimedia files, however, provide the source material that GenAI can transform. GenAI needs large bodies of content to function and, organizationally, that content can exist in unstructured repositories or as unstructured data inside of structured systems.

## 1.2. Content as Competitive Advantage

Fundamentally, the power of GenAI at an enterprise level comes at the intersection of the organization's proprietary information and the GenAI engines. For differentiation and competitive advantage, organizations should focus on their core competencies and "doing what they do best." Thus, content that is unique to an organization should be treated as a high-value and critical asset, and it should be the focus for applying GenAI.

## 1.3. Three Factors for Success

To fully leverage the value of the unstructured data and succeed with GenAI, organizations must address three key factors:

**Employee Engagement**
*Employees must be supported in a way that ensures their safety while exploring opportunities for benefit.*

**Content Access**
*Valuable content is accessible and secure inside corporate repositories and systems.*

**Content Hygiene**
*To improve the relevance and accuracy of GenAI results, data sources must be cleansed and maintained.*

These three factors are inspired by the principles of organizational change management and information management. Based on decades of proven practice, these factors will help you position your organization for success with GenAI.

# 2. Introduction

Artificial intelligence has been a part of the information management industry for decades. Narrow – or purpose directed – AI has powered forms recognition, optical character recognition (OCR), automatic classification, and more for years. The transformative change in AI for organizations comes in the form of GenAI, which can be applied to a wide range of tasks – rather than being solely focused on a specific use case.

The power of GenAI comes from the large amount of content that is used to train the models. This provides the engine with the ability to predict patterns of responses and, ultimately, generate answers based on those predictions.

Additionally, GenAI can leverage additional content post-training to contextualize the responses. Called "Retrieval-Augmented Generation" (RAG), this is what gives GenAI the ability to be adapted to generic situations and customized to organizational standards – written or unwritten.

GenAI providers are developing solutions that allow customers to create contextualized surfaces for these models. By predefining prompts, changing the way that RAG feeds content into the model, and adjusting various parameters, GenAI can be focused into specific use cases.

To accomplish the transformation that leads to use of GenAI for business advantage, organizations must address three key areas: employee engagement, content access, and content hygiene. We'll briefly summarize these critical success factors and will address them more fully throughout the paper.

## 2.1. Employee Engagement

Large scale projects – especially those involving change – fail at a rate of around 70%. This estimate has remained with us for decades, and the research around the rate of failure for change projects doesn't seem to indicate that it's getting better. When asked why projects fail, the most common response is that the users didn't make the change. Whether it's digitization of paper processes that they never adopted or reversion to previous approaches, organizations make large investments and observe flagging results, because the users don't embrace and engage with the change.

Creating the kind of engagement necessary to adopt a transformative technology like GenAI requires organizations to understand how users try and respond to changes. Many models for change are available including the work of Everett Rogers, which is discussed below.

Organizations have specific "levers" they can pull to increase user adoption and engagement. By making things appear simple and making the approaches compatible and safe, users are more likely to engage with the desired approaches. The work of Richard Florida (Florida, 2012) and Edward Deci (Deci & Flaste, 1995) indicates that we can no longer order people to do behaviors that aren't comfortable to them and the way they work. Organizations must focus on motivating and enabling users.

Effective change also means addressing the lack of psychological safety in the organization – at least as it relates to trying and failing with GenAI – so that learning can happen and the organization can ultimately succeed.


## 2.2. Content Access

The key insight that effective organizations have discovered is that the value of GenAI comes from the use of the retrieval-augmented generation (RAG). RAG is a powerful strategy that allows prebuilt large language models (LLMs) to provide context-specific results that match the organization's tone and reflect the unique knowledge of the organization. The impact of the RAG strategy with GenAI is unlocking organizational knowledge assets to create value.

Content access requires sensitivity to permissions and other restrictions, so that the augmentation with the content only happens for users authorized to use the content. This process requires gathering these permissions and restrictions from the various systems in use.

Most organizations will find the data they need in multiple systems. Some unstructured data resides in the structured walls of the core systems and databases. It's highly unlikely that the organization will have only one unstructured repository for files. It's more likely that organizations will need to invest in getting the GenAI tool access to the content stored in other systems.

The continuing process of bridging these gaps and making content available will be the way that new uses for GenAI are unlocked in the organization.

## 2.3. Hygiene

RAG is fundamentally built on top of search retrieval. Only the most relevant results to the user's inquiry can be used to augment the model so the importance of search relevance can't be overstated. One of the consistent ways that we can improve search relevance is through the application of content hygiene. From the elimination of inappropriate language, outdated, and trivial information that may reduce the effectiveness to the organization that makes it easier to associate with topics, hygiene is focused on its impact on search.

When organizations have designed and implemented strong metadata and taxonomy plans, search is more readily able to identify the most relevant items to a query. Metadata fields can be weighted more heavily, thereby elevating the content ranking in search results.

## 2.4. How to Use this Whitepaper

This whitepaper is intended to be shared with the key stakeholders of a GenAI project. The high-level summaries are designed to be read by leadership at any level. The included assessments are designed to be completed by those who are closest to the work. We expect that leaders will share the document – with their comments and direction – with their teams to result in an assessment of an organization's readiness to implement GenAI projects.

The assessment scores are designed to create a mechanism for identifying progress for those organizations who do not appear to be able to successfully implement a GenAI project. The questions themselves reveal specific capabilities, attitudes, and allocations of resources that need to be made to create the best likelihood for success.

We encourage you to use the assessments at the macro level for general readiness and at the micro level to identify key changes that you can make to become more ready.

To get engagement, an organization must foster two critical culture factors:

1. ***Change acceptance***
2. ***Change skills***

At first glance, it would seem that AI adoption is about technology, but nothing could be further from the truth. It takes creative people to apply the technology in ways that generate value. We must first let go of the preconceived notion that we must hire AI experts – that don't exist – to leverage AI in our organizations.

After disavowing ourselves of required technical wizardry, we can focus on the critical need to engage, enable, and motivate our current workforce to try, fail, and, ultimately, succeed with a new way of working.

# 3. Employee Engagement

Over decades of information management projects, the one key distinguishing characteristic of success is human engagement with the system. Large-scale investments in modern technology fail not because of technical limitations or challenges, but because users refuse or lack the training to use the technology in ways that enhance the organization's performance.

## 3.1. You Don't Build It

Before the proliferation of cloud computing, there was a demonstrated need for deep technical resources to implement infrastructure components for the enterprise. With the rise of cloud computing, the need for in-house technical talent has decreased. Organizations no longer need in-depth technical specialists, nor do they need to develop non-proprietary services. These services come from cloud-based providers.

While every organization should maintain technical expertise, that expertise is now more focused on the aspects of technology that users interact with. Technical resources are now focused on solving business problems by assembling resources both internal to the organization and those purchased as services. GenAI services are available from the cloud vendors that you work with, which you need only to contextualize and integrate. The focus for any GenAI readiness project should be on how you can enable your employees to appropriately leverage the technology that you can license for use.

The result of this shift is an overall focus on how you can build solutions to organizational challenges and opportunities rather than focusing on making basic infrastructure available.

### Engagement or Buy-In

*Organizations speak about buy-in to their initiatives. Buy-in simply means they'll comply with the instructions that are given. While this is a good start, it falls short of the level of engagement that is necessary when we're embarking into unknown territory. If we're only achieving buy-in, we're leaving a lot of opportunity untapped.*

*In the 1990s, it was a popular tactic amongst corporations to reset employees' browsers to the corporate intranet or, worse, to the public internet site. Did we get buy-in from employees to use the pages? Objectively, yes; employees did, in fact, visit these pages. However, upon closer inspection of the data, we learned that employees rarely interacted with pages by clicking links and instead went immediately to a toolbar link for wherever they really wanted to go.*

*On this GenAI journey, we need users to want to explore, discover, and leverage GenAI technologies.*

## 3.2. Change Acceptance

Some organizations struggle to make any kind of change, while others seem agile, nimble, and responsive. These organizational-cultural factors can have a profound impact on the organization's ability to leverage GenAI for competitive advantage. The less trust, tolerance for ambiguity, and engagement there is, the more organizations will struggle to implement any kind of change, especially the kind of profound change in the working style that GenAI requires. Some organizations have built in distrust for any kind of change. Changes have become unsafe, because people will become more overworked, will be berated, or may even lose their jobs. The factors here are key to minimizing friction during the transition or even lubricating the change process.

### 3.2.1 Adoption Change Curve

Much of what we know about how to drive adoption comes from the fields of Iowa and the work of Everett Rogers. Among other insights, Rogers came up with the adoption curve (Figure 1), which specifies the well-known innovators, early adopters, early majority, late majority, and laggards categories based on a standard distribution "bell curve" (Rogers, 1962).

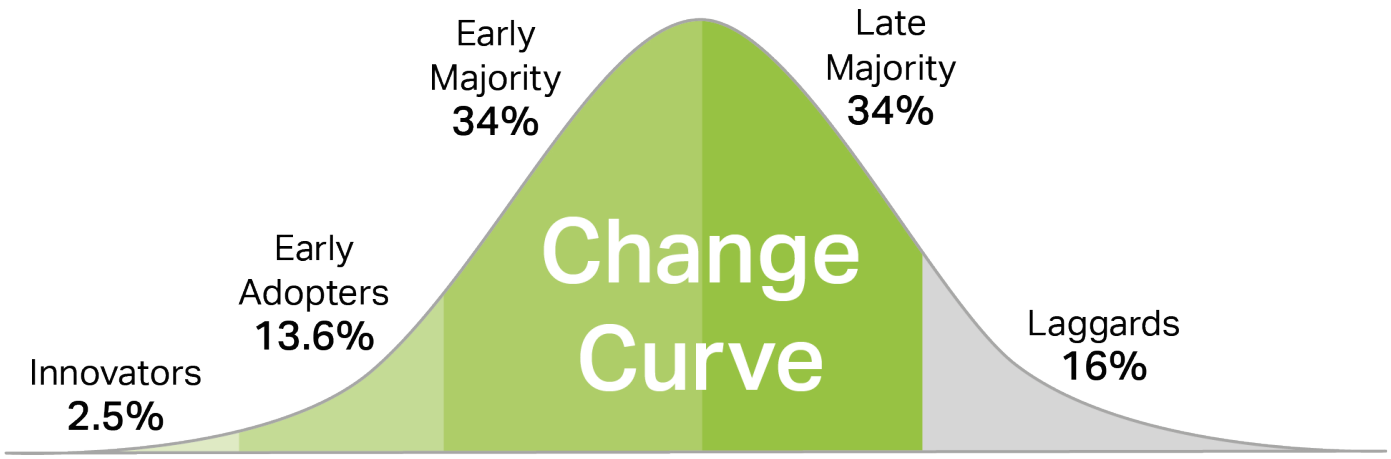



*Figure 1: The Adoption Change Curve*

Rogers was careful to explain that the process of generating large-scale change requires innovators, but the majority of the users (both early and late majority) won't directly listen to or be influenced by innovators, because they're too different. The majority needs evidence of success from the early adopters to make it safe enough for them to try.

Organizations often work to curtail innovators that go outside of approved channels to leverage new and interesting technologies. These innovators are great signals to the organization about what can be done with new technologies including AI – and can signal where governance will be needed to prevent problems. Effectively harnessing innovators to spark the interest of early adopters is an important part of a healthy change readiness.

### 3.2.2 Accelerating Adoption

Rogers' contributions to change management aren't confined to the relatively popular change curve. Rogers also proposed that the most successful innovations leveraged five factors for improving adoption. They are:

- **Relative Advantage –** *The innovation needs to provide an advantage to the person themselves. GenAI does well here, because it has the potential to help the person be more effective.*
- **Compatibility –** *The innovation must be compatible with the existing ways of working. The more GenAI is integrated into existing approaches, applications, and systems, the more it will be adopted.*
- **Simplicity –** *The innovation must appear simple. Here, Rogers' language was against complexity – but as a motivating factor, apparent simplicity is key to adoption. The more that GenAI responds without having to learn prompt engineering, the better.*
- **Trialability –** *The innovation must be trialable. The person must be able to go back to the old way of work. This means that GenAI solutions should be deployed alongside existing ways of working.*
- **Observability –** *The innovation's use must be observable to others in the organization. Celebrating the use of GenAI and sharing that internally will improve its adoption.*

When encouraging change, we should consider Rogers' insights as well as those of more recent scholars.

### 3.2.3 Psychological Safety

Amy Edmondson popularized the concept of psychological safety in her book *The Fearless Organization* (Edmondson, 2019). She discovered that what distinguished teams with high performance from those with low performance was their willingness to try new things and fail. This was driven by the degree to which team members felt as if they could experiment and make mistakes without repercussions. Her subsequent work, *Right Kind of Wrong* (Edmondson, 2023), builds on the necessity of making mistakes and learning from them.

Creating psychological safety allows employees to do the experimentation necessary to take advantage of new technologies. It's impossible to learn from mistakes if you're in constant fear that you'll be terminated, demoted, or otherwise punished for the risk-taking behavior that is needed when learning how to best leverage GenAI for your organization. According to a Bankrate survey, 72% of Americans don't feel financially secure. Of those, 26% believe they'll never be financially secure (Foster, 2023). Asking people who don't feel financially secure to risk their livelihood for some corporate project simply won't happen.

Organizations with lower levels of psychological safety find that it's difficult to get employees to experiment with GenAI and get the process started. They'll often find they have few – or no – innovators, because they've selected them out. Creating a culture where failure is accepted and even expected creates the kind of safety necessary to rapidly engage with GenAI.

## 3.3. Change Skills

The general conditions, like psychological safety, form the baseline from which organizations work towards change. However, there are change management specific skills and competencies that can be used to accelerate engagement. Organizations that want to take advantage of GenAI need to create ways for users to leverage GenAI and need to create reasons for them to try it. Change management skills create the motivation for individuals to engage.

These change management skills are complementary to, but not replaced by, the project management and operational excellence skills in the organization.

Richard Florida, in his book, *Rise of the Creative Class*, explains the shift in job classes that has changed the way that organizations must motivate staff (Florida, 2012). He explains that today's workers do heuristic rather than formulaic jobs. The result is that we can't order people to do something they don't want to do. With formulaic jobs, it's relatively easy to both force compliance and monitor for non-compliance. However, when there's no right way to build the next presentation, it's hard to say whether they're doing it using the GenAI tools that you want them to use or not.

Edward Deci and Richard Flaste, in *Why We Do What We Do,* express a similar perspective differently. They explain that intrinsic drive comes from autonomy, mastery, and purpose. We can't just tell people they must use AI, because it will negatively impact their intrinsic drive (Deci & Flaste, 1995). We'll necessarily be impinging on their autonomy. As a result, we need to take a motivational rather than dictatorial strategy.

### Machine or Plant

*The industrial revolution was a powerful step forward in the history of humanity. The side effect of the industrial revolution is that we began to see everything as if it were a machine, a set of replaceable components. The problem with this mentality, as every leader has learned, is that people are not replaceable components. The machine view of the organization allows for efficient breakdown of tasks, but it comes at the cost of emergence.*

*One of the previous models of thinking was agricultural, where the goal was to create the right conditions for growth. Ensuring the right soil, water, and sunshine allowed crops to grow without the precision of a machine. We could still plant seeds in straight lines at equidistant spacing, but more was left up to mother nature. When enabling GenAI in your organization, better results may be received by concentrating on enabling conditions and removing barriers instead of machine-like planning.*

*When we're talking about creating change and motivating engagement, the machine model doesn't work. As difficult as it may be to accept from a command-and-control point of view, when approaching change, we must accept that we can't always command the behavior we desire. Instead, in some cases, we must create the conditions that allow for the growth we want to see.*

*Readers interested in different views of organizations are invited to read Images of Organization (Morgan, 2006).*

**Change managers need the skills to:**

- *Identify the specific behavior changes necessary to take advantage of GenAI.*
- *Create meaningful personas of personnel in the organization to target.*
- *Create compelling visuals that generate both understanding and desire.*
- *Motivate personnel to want to make the change for themselves as well as for the organization.*
- *Develop a communication strategy that clearly communicates the objectives and milestones for success as well as the vision.*
- *Motivate all corporate communicators to improve their efficacy when communicating about change.*
- *Build strong stakeholder teams and coalitions to remove barriers and encourage participation.*
- *Identify high-leverage changes that can lead to big results with minimal resistance.*

A desire for change at a senior level is important but insufficient to drive change across the organization. Skilled change managers remove barriers and enable success.

## 3.4. Personnel and Engagement Assessment

Complete the following questions to identify opportunities for improvement in your current organizational readiness and capabilities to engage your organization's employees in the change process needed to effectively implement GenAI in your organization.

You can complete this assessment online by going to https://info.aiim.org/personnel-and-engagement-assessment

### 3.4.1. Organizational Culture

**Rate the following statements in terms of your certainty.**

| Statement | Very uncertain | Somewhat uncertain | Ambivalent | Somewhat certain | Very certain |
|---|---|---|---|---|---|
| How clearly can you define your organization's strengths and weaknesses in terms of leveraging AI? | ☐ | ☐ | ☐ | ☐ | ☐ |
| How clearly do you feel you can articulate the external opportunities and threats to your AI initiative? | ☐ | ☐ | ☐ | ☐ | ☐ |

**Rate the following statement in terms of how strongly you believe in it.**

| Statement | Firmly believe | Slightly believe | Undecided | Slightly disbelieve | Firmly disbelieve |
|---|---|---|---|---|---|
| To what degree do you believe that your AI projects should be implemented with strong, preplanned, "waterfall" methodologies? | ☐ | ☐ | ☐ | ☐ | ☐ |

### 3.4.2. Organizational Change Skills Assessment

**Rate the following statements in terms of frequency.**

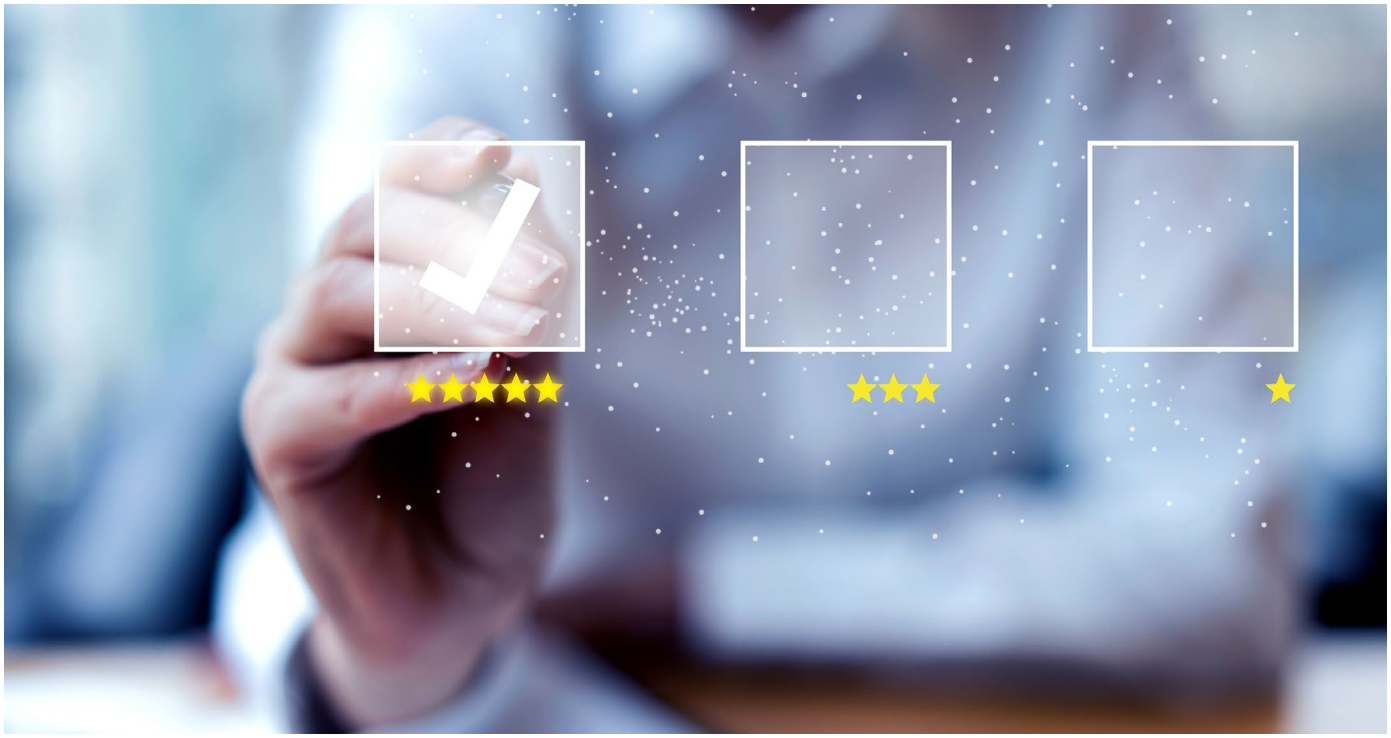
| Statement | Rarely or Never | Infrequently | Sometimes | Frequently | Usually or Always |
|---|---|---|---|---|---|
| When establishing metrics for your changes, how often do you feel you delineated predictive leading measures and rear-view lagging measures of success to evaluate your progress? | ☐ | ☐ | ☐ | ☐ | ☐ |
| How often do you use personas when communicating about change projects? | ☐ | ☐ | ☐ | ☐ | ☐ |
| How often do you use visuals when communicating about change projects? | ☐ | ☐ | ☐ | ☐ | ☐ |
| How often do you utilize a standardized method to assign roles and responsibilities (e.g., RACI charts) when organizing your change projects? | ☐ | ☐ | ☐ | ☐ | ☐ |
| When analyzing your stakeholders, how often do you assess each stakeholder's motivators? | ☐ | ☐ | ☐ | ☐ | ☐ |
| When leading change initiatives, how often do you analyze which hidden barriers, like cognitive biases, affect both you and your organization? | ☐ | ☐ | ☐ | ☐ | ☐ |
| When communicating key messages about change initiatives, how often do you create different messages tailored to each persona? | ☐ | ☐ | ☐ | ☐ | ☐ |
| When you lead change initiatives, how often do you create a dedicated communications plan that includes timing, purpose, sender, recipient(s), media channel used, level of detail, and key messaging? | ☐ | ☐ | ☐ | ☐ | ☐ |

| Statement | Rarely or Never | Infrequently | Sometimes | Frequently | Usually or Always |
|---|---|---|---|---|---|
| When you write announcements about change initiatives, how often do you include all of the following: who initiates the change, who is impacted, what is changing, when it is changing, where the impacts of the change will be felt, why the change is being initiated, and how the change process will work? | ☐ | ☐ | ☐ | ☐ | ☐ |
| When you communicate about change initiatives, how often did you consider the cognitive load on the reader? | ☐ | ☐ | ☐ | ☐ | ☐ |
| When categorizing stakeholders, how often did you quantify each stakeholder's salience in the three dimensions of power, legitimacy, and urgency? | ☐ | ☐ | ☐ | ☐ | ☐ |
| How often do you evaluate stakeholder motivators afrom both personality and positional perspectives? | ☐ | ☐ | ☐ | ☐ | ☐ |
| When managing change initiatives, how often do you establish standards for expected behavior and accountability? | ☐ | ☐ | ☐ | ☐ | ☐ |
| During change initiatives, how often do you establish a system to address poor behaviors, including what verbiage is used to signal behavior violations? | ☐ | ☐ | ☐ | ☐ | ☐ |
| When closing out change management initiatives, how often do you perform an after-action review? | ☐ | ☐ | ☐ | ☐ | ☐ |

### 3.4.3. Psychological Safety

**Rate the following statements in terms of your confidence in your abilities.**

| Statement | Very unconfident | Somewhat unconfident | Unsure | Somewhat confident | Very confident |
|---|---|---|---|---|---|
| How would you rate your ability to analyze whether someone's stated requests are really what they want or just the means to a different end? | ☐ | ☐ | ☐ | ☐ | ☐ |
| How effectively can you manage stakeholder conflict? | ☐ | ☐ | ☐ | ☐ | ☐ |
| How effective are you at managing the perception of fear in your organization? | ☐ | ☐ | ☐ | ☐ | ☐ |



The next few questions ask you to review a set of techniques or tools you've used. For each question, mark the appropriate options, or write in your own answers in the space provided.

| | | |
|---|---|---|
| **1. The last time you led a change initiative, which of the following types of factors did you analyze for their impact on the initiative?** | ☐ Political<br>☐ Economic<br>☐ Social<br>☐ Technological<br>☐ Legal<br>☐ Environmental | ☐ Other |
| **2. The last time you performed a risk assessment, which of the following strategies did you employ to manage the risks?** | ☐ Avoidance<br>☐ Monitoring and Preparation<br>☐ Mitigation<br>☐ Transference<br>☐ Acceptance | ☐ Other |
| **3. The last time you established metrics for your change, which of the following categories did these metrics fall under?** | ☐ Financial<br>☐ Internal<br>☐ Customer<br>☐ Innovation and Learning | ☐ Other |
| **4. Which of these factors, according to Everett Rogers, accelerate adoption?** | ☐ Relative Advantage<br>☐ Compatibility<br>☐ Apparent Simplicity<br>☐ Trialability<br>☐ Observability<br>☐ Knowledge<br>☐ Attitude<br>☐ Practices | ☐ Other |

| | | |
|---|---|---|
| **5. In your last change project, which of the following educational approaches did you use?** | ☐ Instructor-Led Training<br>☐ Computer-Based Training<br>☐ On-the-Job Training<br>☐ Planners<br>☐ Sidekicks | ☐ Other |
| **6. The last time you planned communications about a change initiative, how many channels did you plan for? Select only one option.** | ☐ 1<br>☐ 2<br>☐ 3<br>☐ 4<br>☐ 5 or more | |
| **7. The last time you planned communications about a change initiative, which criteria did you use to determine what channel to use?** | ☐ Urgency<br>☐ Direction<br>☐ Scale<br>☐ Purpose<br>☐ Audience<br>☐ Length<br>☐ Findability | ☐ Other |

### 3.4.4. Scoring

To develop an overall score, start by tallying the scores for the first 21 questions, from 0 for answers in the leftmost column to 4 for answers in the rightmost column. For the final seven questions, tally a 0 if you marked less than one third items, a 2 for one third to two thirds, and a 4 for more than two-thirds of the items.

For the final seven questions, use the rubric below to score your answers based on the number of options you marked. For question 6, score your answers based on the single option you selected.

| Question | 0 Points | 2 Points | 4 Points | Your Score |
|---|---|---|---|---|
| 1 | 0-2 | 3-4 | 5+ | |
| 2 | 0-2 | 3-4 | 5+ | |
| 3 | 0-1 | 2-3 | 4+ | |
| 4 | 0-2 | 3-5 | 6+ | |
| 5 | 0-2 | 3-4 | 5+ | |
| 6 | "1" or "2" selected | "3" or "4" selected | "5 or more" selected | |
| 7 | 0-2 | 3-4 | 5+ | |

When you have your total number, divide it by 1.12 to receive your score. (You're dividing by the 28 questions and then multiplying by 25, which adjusts the number to a 100-point scale.) This number is your personnel engagement score.

If you fill out the online assessment at https://info.aiim.org/personnel-and-engagement-assessment we'll do this scoring for you.

# 4. Content Access

The rise of search engines in the 1990s transformed the internet from a cryptic place of connecting to those things you knew to a place where you could find anything. The rise of enterprise search in the 2000s promised similar advances for the organization. At the time, these efforts were widely denounced for failing to provide the promised impact. However, in the two decades that have passed since then, the way of work has adapted from one of navigating to one of predominantly searching for content. Additionally, improvements in relevancy have enabled greater findability.

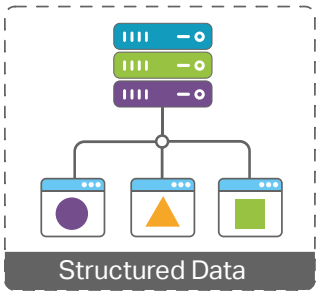
Search and GenAI fundamentally need the same things to be effective. They need reliable access to the information in the organization as well as its related restrictions. Those restrictions might include direct permissions, sensitivity, and digital loss protection rules. If your organization has a robust enterprise search strategy that enables your search engine to gain access to content across the important systems in your organization, you have the groundwork laid to supply the GenAI engine the information it needs. Sadly, few organizations have robust enterprise search strategies.

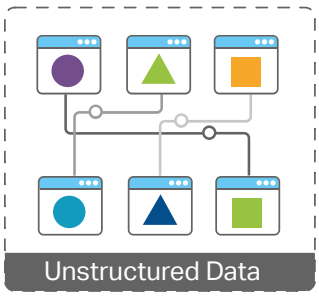
This section describes the aspects of a content access strategy that are applicable to both search and GenAI.

## 4.1. Structured and Unstructured

While the benefits of GenAI have mostly come from unstructured data, much of the unstructured data still resides in structured repositories. Structured repositories are those that store rows and columns of information, like databases and enterprise systems. Unstructured data is most frequently found in documents, presentations, emails, and notes that do not follow a rigid structure.

Where these concepts overlap is in the text fields of the structured repositories, where there may be large volumes of text information. Consider the system supporting a call center, which includes notes written by every agent on every call. This is unstructured data stored in the structured repository of the call center management system.




When we're talking about content access for GenAI, we need to consider the information stored both in the traditional, unstructured repositories but also in those structured repositories that contain significant free-text entry.

## 4.2. Restrictions

When establishing content access, there is an often overlooked need to maintain restrictions. The restrictions associated with the information are critical to capture to prevent unintended use and disclosure of sensitive information. In practice, gathering the content itself and addressing the restrictions must be addressed simultaneously. For explanation's sake, in this section, we'll focus only on the restrictions on the content. We discuss the content itself in the next section.

For a GenAI system to be trustworthy, it must not disclose information that the user shouldn't have access to. This means that, before using material to generate a response or providing the material directly to the user, it must ensure that they're entitled to see it.

This access control process relies on three restriction components, the first two of which are essential for every repository:

1. **Identity –** *The identity of the users who can access the content must be discoverable.*
2. **Permissions –** *The rights that individuals and groups have to access the content must be discoverable.*
3. **Sensitivity –** *The sensitivity of the item must be knowable. For instance, inclusion of personally identifiable information (PII) may be further restricted even when permissions are present.*

We start with a discussion of the implications of pre- and post-processing the restrictions before examining each of the essential restriction components. To understand the reasons for pre- and post-processing of restrictions, it's necessary to understand identity, permissions, and the relationship to the data.

### 4.2.1. Identity

The ability to unify identity across repositories is an often-overlooked challenge for establishing content access. While a single identity can be used for the GenAI system to index and access the content, ultimately, the permissions and restrictions must be encoded in a way that's consistent across all the repositories. Legacy repositories may be using system-defined authentication with usernames and passwords. These present challenges, as the AI system uses a unified network identity.

Modern systems use LDAP, Kerberos, or claims-based authentication strategies that ensure that the login for the content repository is consistent with the user's network credentials, making the mapping of identities transparent. If your environment includes legacy systems without integrated identity management, you may need to build an identity mapping mechanism.

### 4.2.2. Permissions

The most common restriction, by far, is the permission applied to the item – or column – to a set of users and groups. Luckily, permissions are largely assigned to groups and users to access a container and all its items. These cases are well supported across search and GenAI platforms. The repository simply transmits the list of users and groups and their access to the search or GenAI engine for later filtering.

While most repositories maintain a rich set of access rules, search and GenAI engines, in most cases, simplify access to a simple read or no-read access owing to their narrow use case for the content. While the complexity of multiple permissions is removed, the need to understand the shape and scope of the read permission granted can be challenging.

#### *4.2.2.1. Column Level*

For the most part, search and GenAI tools are unable to utilize column-level data. For instance, perhaps only managers can see an employee's salary. While column-level access to data is useful in some cases, it requires some unique strategies when working with search and AI.

The most common strategy is a splitting strategy, where the content appears in the repository as two separate entities. The first entity includes only the content without column level security. The second entity includes the column-level-restricted content and a connection to the rest of the record.

In general, these cases are not well addressed by in the tools available today; other than the splitting strategy above, it's recommended that these use cases be deprioritized where possible.

#### *4.2.2.2. Relational*

Far more complicated permissions are those based on the relationship between the person and the data being accessed. For example, imagine that you're a provider in a large health system attempting to access a health record. It may be that the system restricts access to those patients who've asked for services or whom you've seen in the past. Thus, determining whether a patient record is accessible or not requires reviewing the user's current elections as well as reviewing the patient history for an entry with your name.

In some cases, potentially sensitive information, like personal health information, may need to be excluded from the results of search of GenAI even when the user does have rightful access.

These sorts of relationships require either projecting additional access control entries for each possible combination or evaluating on a post-processing basis. It's these relational conditions that are the best use cases for post-processing of restrictions.

#### *4.2.2.3. Pre and Post Processing*

Every search and GenAI tool provides a mechanism to deliver the restrictions to the tool for its own filtering. This is referred to as pre-processing of the restrictions, because they're embedded into the tool's filtering process. Some tools support a post-processing approach, which allows the repository – or, more frequently, an interface layer – to respond to the tool wanting to know whether a user has or doesn't have access to a set of content.

Post-processing is important when the restrictions don't fit the narrowly defined confines of typical user or group to content restrictions. Post-processing allows for richer permission scenarios with the caveat that the post-processing service is in-line with the customer response; therefore, it must be highly performant and scalable. Post-processing, while enabling enhanced scenarios, should be used as minimally as possible.

### 4.2.3. Sensitivity

While content permissions represent the most prevalent restrictions on the availability of content, there are other conditions that may apply to content that would further restrict access. These are often described as sensitivity. Sensitivity ratings of content may constrain where information is accessible as well as additional requirements for access based on corporate policy.

Consider trade secrets, which allow an organization to be competitive. These should be available via search and to GenAI but also should be restricted in ways. It may be that the content can only be accessed while on-site or require explicit reauthentication. These sensitivity aspects of the content should be conveyed to the index where possible – and if not possible, post-processing of access should be addressed.

One common case for sensitivity is PII, which can be programmatically assessed through tools and features. This is frequently called data loss protection (DLP). In these scenarios, the DLP engine discovers a pattern of characters that match important information that requires additional sensitivity. This might include particular types of document identifiers but more frequently refers to PII markers. A social security number in the United States or social insurance number in Canada would cause the system to identify the document as containing PII.

Inherent in DLP is the concept that a single occurrence is categorically different than multiple occurrences. For instance, a single credit card number isn't ideal. It's certainly something that should generate a warning. However, a spreadsheet containing every corporate card in the company conveys a different level of response.

Ultimately, sensitivity is a broad category of non-permission restrictions that must be accounted for when making content available to AI engines.

## 4.3. Content Itself

The permissions related to content aren't useful without the content itself. Ultimately, this content needs to be indexed from repositories and prepared into a sequence of content that the search or AI engine can process.

### 4.3.1. Repositories

For each of the content repositories in the organization where access is required, an interface for the search or Gen AI engine is necessary. According to the AIIM 2023 State of the Intelligent Information Management Industry Report (Association for Intelligent Information Management), the number of organizations that have 7-10 discrete information systems or repositories has doubled. Organizations of moderate to large size are almost certain to have multiple repositories where their content is stored. For generative AI to leverage this content, it must be accessible.

The same report indicates that 74% of these systems are not connected to other line of business systems, indicating that, historically, little work has been done to integrate these systems – critical work that needs to be done to take advantage of GenAI.

Most modern repositories will support a set of web service (SOAP) and REST endpoints, which can be used for identifying and retrieving the content in the repository. Older systems may require COM-based interfaces, or even custom code to extract the information from underlying databases and file storage outside of system itself. Third parties also offer interfaces to many repositories to connect them with the generative AI platform of your choice.

Repository access in whatever form may need to be rate-limited by the connector to balance the needs of indexing the content quickly and placing too much load on the system.

### Not All AI is the Same

*Information management tool vendors have been and continue to integrate AI tools into their offerings. Optical character recognition (OCR) has been a staple AI tool for converting images to text for decades. Other AI-based tools like auto classification are also used to augment information in these systems. However, these claims don't mean that your organization can leverage your proprietary content as a part of generative AI unless that data is made available to a generative AI engine.*

*Because organizations are faced with multiple systems, and the power of generative AI is making all this information available to one engine, individual repository GenAI tools will be of limited use in the larger organizational advantage context.*

#### *4.3.1.1. Indexing Load and Time*

There are two conditions for the indexing of a content repository. There's the full index condition, where all content is indexed, and there's an incremental index condition, where only changed content is indexed. These conditions have different impacts on the repository itself.

Full indexes of the content are required when a repository is added and whenever index corruption indicates that the repository's representation is or may be no longer valid. Full indexes are important, because they can generate a substantial load on the repository, which may interfere with user responsiveness from the system. Also, owing to potentially large amounts of data, these indexes may take a substantial amount of time. During the indexing process, the full extent of the repository won't be available, and therefore the overall length of time to index a repository can mean a substantial delay in content accessibility.

As a result, it's important to assess both the impact on a repository to perform an index and the total time necessary for the full index to be completed. In practice, this is a balance that is initially set and adjusted over time.

Incremental indexes can either be triggered to happen continuously as content is added or changed in the repository or, more commonly, added through a polling process that asks the repository for items changed since a last crawl token. The polling process timing needs to be balanced between creating too much load on the repository and delaying availability of the data for too long.

### 4.3.2. Content Pipeline

Unstructured data has a variety of different markups and encoding. The format for a Word document file and an AutoCAD file are radically different. Search and GenAI platforms need the extracted text from these files for processing. That means the indexing process is processing the file such that the result is a string of data.

In the case of search (and, by extension, GenAI), the content return includes some weighting information that conveys the importance of explicit metadata as well as in some inferred content. For instance, the largest text on the first slide of a presentation is inferred to be the title of the presentation and is therefore weighted more strongly.

If you have custom formats for files or markup that is substantially different than standard, it may be necessary to add a content processor to help the search or GenAI engine to properly consume the content. Content processors are tools designed to convert file formats or sources into a stream of text that can be ingested. For instance, in the Microsoft Windows ecosystem, IFilters are used to convert Microsoft Office documents, PDFs, and other file types into text streams. Third parties offer a wide range of content processors for extracting content from less frequently used file types.

## 4.4. Relationships

In some cases, most notably Microsoft 365, search queries and the GenAI responses are further enhanced for relevance and disambiguated through awareness of relationships. Microsoft Graph is an automatically generated social network analysis of the relationships between people. Social network graphs contain nodes (people) and edges (relationships). The edges are formed by processing a set of signals that indicate there may or may not be a relationship between people. Edges have types and strengths which convey what kind of a relationship it is and the strength of that relationship. Generally, more signals equate to stronger relationships. The process of sending signals to the AI engine allows it to build more accurate relationships and thereby better disambiguate when the query itself isn't clear.

A very strong signal sent to the graph can be directly converted to an edge. For Instance, you could send in the corporate organizational structure from an HR system. Most of the time, signals are very lightweight interactions that lead to larger meaning. Sometimes, the signals can be stronger, but not in the core information systems. For instance, imagine that your manufacturing floor is made up of manufacturing cells with groups of four people. These four people work together to create products. However, the only place where this relationship is stored is in the manufacturing execution system. Sending the signal to the social network graph about the degree to which people work together on a daily basis can help AI return results that are generated from your peers as most likely sources of the answers you're looking for.

Consider an interaction with GenAI about a problem where a response prompts a new query, "What would Rob do?" The GenAI engine must first determine who Rob is. Querying the social network graph, it can determine that there are three Robs in your close network: your immediate manager, your peer, and your subordinate. It can ask you for clarification about which Rob you mean and provide you a response with a list to choose from. In cases where you only have one Rob in your social network graph, it may assume you mean that Rob, since it's the most probable answer.

## 4.5. The Content Access Assessment

The content access assessment is different from the other two assessments included here. It is a matrix consisting of the repositories in one dimension and a set of six key criteria in the other dimension.

### 4.5.1. Repository List

To assess content access, we start by identifying the content repositories in the organization, including those holding both structured and unstructured data. This list may already be partially or completely available as an application list.

For each of these repositories, there are six questions that will assess the ability to leverage the content in those repositories for GenAI.

### 4.5.2. Repository Type

Repositories store different kinds of data. When evaluating the repository type, consider the following options:

1. **Structured data without unstructured data –** *Databases of raw numbers generated by devices often fall into this category.*
2. **Structured data with unstructured data –** *These systems are structured in their design but allow for unstructured data areas. Consider applications like call center tracking and relationship management systems, which have large amounts of unstructured technical data inside the context of their structure.*
3. **Unstructured data without metadata –** *These systems contain documents, but there's no additional metadata assigned. The prototypical example is a file share.*
4. **Unstructured data with metadata –** *These repositories hold files but have associated metadata with them Microsoft SharePoint, IBM FileNET, Hyland OnBase, and OpenText Documentum are all examples of these types of repository.*

If you find this matrix is easier to fill out electronically using a workbook, go to https://info.aiim.org/ai-assessment-excel-spreadsheet

### 4.5.3. Repository Scope

Some repositories have limited scope within an organization, while others have a much larger impact. Generally, the scope can be divided into:

1. **Narrow –** *Utilized in or benefiting a single team or department.*
2. **Moderate –** *Utilized in or benefiting a department or division.*
3. **Large –** *Utilized in or benefiting multiple departments or divisions.*
4. **Global –** *Utilized in or benefiting the entire* organization.

### 4.5.4. Content Utility

The most subjective of the assessment criteria, the utility is a rough guess at the usefulness of the content in the repository for GenAI. While this assessment will be ultimately made on a use-case-by-use-case basis, here, the intent is to capture an overall sense of utility broken into four categories:

1. **None –** *There's no expectation that the content would be of use. Consider scanned historical packing lists as an example.*
2. **Little –** *There is little indication of utility, but there's also little indication that the content will be of high value. Social media posts may be in this category.*
3. **Some –** *There's an expectation of some value to be created by using GenAI, but this value isn't exactly clear. For instance, this might include a library of research papers or the organization's policies.*
4. **A lot –** *The pathway to utility for the content is clear and powerful. For instance, reviewing knowledge bases and curated repositories of content and making it available through GenAI.*

### 4.5.5. Repository Volume

It's impossible to scale an absolute amount of content in a meaningful way across industries and organizations. However, it is possible to assess the relative amount of content in each repository as a percentage of the organization's total content corpus. Here, some categories are useful:

1. **< 5% of total volume –** *These repositories don't contain much meaningful data. Consider the single-department, proprietary application that only handles a small portion of their needs.*
2. **5-9% of total volume –** *These repositories, while still relatively small, may have sufficiently targeted information to be of real value. These repositories are typically used to coordinate work between a few departments.*
3. **10-30% of total volume –** *These repositories are rarely the largest, but they're sufficiently large to be valuable to the organization. These repositories may include things like a logistics management system, which tracks how orders are packed and shipped to customers and line of business applications.*
4. **> 30% of total volume –** *These repositories are the largest and therefore have the highest potential to generate benefits. One example of this is the enterprise resource planning (ERP) systems of manufacturing organizations. Other industries refer to these sorts of repositories as "core" systems.*

### 4.5.6. Integration and Extensibility

An important consideration for utilization in a GenAI implementation is the degree of difficulty in integrating the system, which is broadly measured in four categories:

1. **None –** *The system includes no known provisions for integration with other systems.*
2. **Documented –** *There are documented integration options from the repository to other systems.*
3. **Third-party available –** *Repositories that are sufficiently popular in the market to have third parties integrating to them already. Presumably, you could acquire an extension, add-in, or adapter that would allow you to connect the GenAI engine.*
4. **Implemented –** *Reserve this option for those repositories that you've already implemented. Repositories with an integration with another system. The capabilities should be like those that would be needed for a GenAI implementation.*

### 4.5.7. Identity Management

The final category of assessment is the degree of difficulty in managing the permissions of the content, measured by categorizing the way that identity is managed:

1. **Proprietary without matching usernames –** *These systems typically use a database-stored username and password system, but, more importantly, the names in this system are inconsistent with the standards used in the core systems. As a result, mapping of users may represent significant challenge.*
2. **Proprietary with matching usernames –** *Authentication is handled internal to the system with a username and password combination, but the username is a one-to-one match to the username used for core identity. The result is an anticipated low degree of mapping complexity.*
3. **LDAP or Kerberos support –** *These systems are capable of legacy authentication through Kerberos and are integrated into the core authentication approach.*
4. **Claims or OpenID support –** *These modern systems can connect with other systems using current identity integration standards and therefore represent the least friction to integrating to a GenAI engine.*

### 4.5.8. The Matrix

| Repository Name | Owner | Type | Scope | Utility | Volume | Integration and Extensibility | Identity Management | Score |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |
| | | | | | | | | |

You can complete the matrix above and score it, or you can get an Excel workbook template that does the scoring for you at https://info.aiim.org/ai-assessment-excel-spreadsheet

### 4.5.9. Scoring

To score by hand, score each repository individually. To do this, for each of the factors except for volume, add 0 points for the first choice, 1 for the second choice, 2 for the third choice, and 3 for the third choice. Multiply the sum by 100 and divide by 15. The resulting number is your score for that repository.

After scoring all the repositories, add all the scores together and divide by the number of repositories – resulting in an average score. Add together the volume percentages for all repositories to get your total volume. If the total volume is less than 90%, multiply the average score by this percentage to get your final score; if the total volume is greater than 90%, your average score is your final score.

# 5. Content Hygiene

Throughout the history of information management, we've had experts whose role was to carefully classify and organize information. Their skills were indispensable for creating indexes of where to find resources. With the advent of search, the internet, and other enabling technologies, many of these professionals found their core strengths no longer essential. Decades have passed since organizations had corporate librarians and a team of information scientists dedicated to helping the rest of the organization find the information they need.

While many organizations decided to forgo their corporate librarians and leave the searching process to the organization's personnel, savvy organizations realized that the techniques that were once essential were now powerful ways to get even more out of the technology that was supposed to have made the roles obsolete.

Content hygiene is the process of ensuring that the content you have available for generative AI meets the organizational needs in terms of appropriateness and consistency and in ways that allow search and GenAI to better identify the most relevant content. It starts with addressing inappropriate language and continues with the common information management tasks of reducing redundant, outdated, and trivial content (ROT), developing taxonomies, managing metadata, and developing a data catalog.

This is the last part of the GenAI readiness assessment, because it's the least impactful. While it still can mean the difference between success and failure, addressing engagement and ensuring content access are more impactful to the success of a GenAI project. Historically information management professionals have been focused on these activities and are largely familiar with performing these activities.

## 5.1. Inappropriate Language

Unless your organization breeds dogs, the word "bitch" in your content is unlikely to be an appropriate word. This and numerous other potentially offensive words are routinely found in the content of organizations. For instance, some information may describe meetings as "bitch sessions." Appropriately removing inappropriate words helps to ensure that they don't inadvertently end up in generative AI output.

More broadly, as the organization makes more content available to the GenAI engine, it's critical that the organization's tone not be disrupted by offensive words that may be stored in content repositories without anyone's awareness.

Regular work to identify and categorize offensive words is necessary to ensure that the content the GenAI engine is trained on won't accidentally spit out offensive results to employees or customers.

## 5.2. Redundant, Outdated, and Trivial

For decades, information managers have struggled with the work to remove ROT from the network to make room for higher value content and prevent the pollution of the search index. In the context of GenAI, these are still appropriate goals, but they are substantially less impactful or concerning than in years past.

Search, and therefore Gen AI, algorithms have learned to respond with the most recent or largest of the redundant content with the expectation that it's the content the user is most likely looking for. Redundant copies do take space, processing time, and attention, but the degree to which they're hidden behind a wall of search similarities makes it less critical to remove them. In most organizations, the largest single cost is personnel, so saving a few dollars on storage costs may not be worth it.

Outdated information, particularly the kind that is clearly wrong, should be removed. However, given the prioritization for more recent information, it's unlikely that GenAI would use an outdated piece of information. Certainly, outdated records should be destroyed in line with their destruction schedules. However, for most organizations, a substantial majority of the content is not identified as records and no formal disposition schedule has been set.

Trivial content has always been the hardest piece of ROT to define and remove, because what is trivial to one person may be vital to another. Because of the enhanced scalability and the greater relevance, trivial content may not be as important to remove as it once was.

## 5.3. Taxonomy

We all need to simplify our world so we can understand and interact with it. Experts, like Gary Klein in *Sources of Power*, explain that we run simulations in our head that can contain approximately three factors and about six states (Klein, 1998). Each of these factors and states can themselves be a reference to another model. The ability to leverage one model inside of another is at the heart of our desire to create hierarchies. A taxonomy is a classification system, often using hierarchies to help us organize and categorize information.

Consistent taxonomic models allow GenAI models to respond better to conversations. The more stable and predictable the organization of information, the more predictable the responses. The skills necessary to develop taxonomies are specialized. Often, organizations find that it is more cost-effective to outsource taxonomic work instead of hiring it, because it's a difficult skill to acquire and is used relatively infrequently. This combination makes contracting these services ideal.

Taxonomies are built on the foundation of metadata, which is often a quicker place to start for organizations looking to make a quick impact on their readiness.

## 5.4. Metadata

Metadata – data about data – is the structured information that accompanies an unstructured document that helps to organize and classify it. Some metadata is intrinsically produced, like the time a file was created or modified. Other metadata must be set by humans or through other automatic classification means. To be able to apply metadata to an item, the fields must first be defined.

Establishing the right number and right types of metadata to ensure efficient categorization and retrieval isn't as easy as it seems. If too much metadata is asked for, it's a burden to set. If too little – or the wrong type of – metadata is allowed for, then documents will be hard to locate, index, and integrate with the GenAI system response.

Having clear guidance on standard and additional metadata fields is important for both consistency and completeness. Consistent naming of metadata elements allows them to be aggregated and assimilated across systems. Failing to allow for the ad-hoc definition of additional metadata prevents potentially useful information from being captured.

Metadata, much like taxonomies, can inform GenAI models and create a greater likelihood that the results received will be valuable.

## 5.5. Data Catalog

Knowing the types and names of metadata attached to unstructured data is a start, but this leaves open the kinds of values in the fields and the meaning of those values. A data catalog specifies not only the fields of data but the value types, ranges, and limits for the fields themselves. This greater degree of precision in the data that is being stored ensures greater predictability of the data and, ultimately, better GenAI-generated results.

## 5.6. The Hygiene Audit

The final portion of the organizational readiness for generative AI is the assessment of practices related to content hygiene. This assessment measures the organization's commitment towards managing the hygiene of its information assets.

You can complete this assessment online by going to https://info.aiim.org/ai-assessment-hygiene-audit

### 5.6.1. Structure and Process

**Rate your agreement with the following statements about structure and process.**

| Statement | Strongly disagree | Slightly disagree | Neither agree nor disagree | Slightly agree | Strongly agree |
|---|---|---|---|---|---|
| The organization has a dedicated function, role, or allocated part of role with responsibility for maintaining data hygiene. | ☐ | ☐ | ☐ | ☐ | ☐ |
| The organization has defined and agreed upon a set of standards for data hygiene. | ☐ | ☐ | ☐ | ☐ | ☐ |
| The data hygiene standards include standardization of metadata and taxonomy as a part of the standards. | ☐ | ☐ | ☐ | ☐ | ☐ |
| There are defined approaches for assessing data hygiene across systems. | ☐ | ☐ | ☐ | ☐ | ☐ |
| The data hygiene assessments are done on a periodic basis (not greater than every 2 years.) | ☐ | ☐ | ☐ | ☐ | ☐ |
| Appropriate taxonomies are implemented in IT systems. | ☐ | ☐ | ☐ | ☐ | ☐ |
| A unified set of taxonomies are implemented across IT systems. | ☐ | ☐ | ☐ | ☐ | ☐ |
| Taxonomies are reviewed on a periodic basis (not greater than every 2 years). | ☐ | ☐ | ☐ | ☐ | ☐ |
| A data catalog has been developed and is maintained (not greater than every 2 years) of all data elements and their sources. | ☐ | ☐ | ☐ | ☐ | ☐ |

### 5.6.2. Accountability

**Rate your agreement with the following statements about accountability.**

| Statement | Strongly disagree | Slightly disagree | Neither agree nor disagree | Slightly agree | Strongly agree |
|---|---|---|---|---|---|
| The organization has defined and agreed upon a set of operational targets for the data hygiene function. | ☐ | ☐ | ☐ | ☐ | ☐ |
| The data hygiene function has appropriate skills to achieve operational targets (including but not limited to data analysis, taxonomy, and data catalog generation skills). | ☐ | ☐ | ☐ | ☐ | ☐ |
| The data hygiene function has sufficient resources to achieve operational targets. | ☐ | ☐ | ☐ | ☐ | ☐ |
| Leaders can adhere to data hygiene standards (over the long-term). | ☐ | ☐ | ☐ | ☐ | ☐ |

### 5.6.3. Integration

**Rate your agreement with the following statements about integration.**

| Statement | Strongly disagree | Slightly disagree | Neither agree nor disagree | Slightly agree | Strongly agree |
|---|---|---|---|---|---|
| Data hygiene is a part of system selection criteria and system development design. | ☐ | ☐ | ☐ | ☐ | ☐ |
| Automation (including but not limited to automatic classification, and workflows) is used to augment unstructured data with additional metadata. | ☐ | ☐ | ☐ | ☐ | ☐ |
| Automation (including but not limited to choice lists, intelligent defaults, and templates) are used to ease encourage and enable user entry of metadata. | ☐ | ☐ | ☐ | ☐ | ☐ |

### 5.6.4. Scoring

When you're finished, tally a score for each of the statements above, from 0 in the leftmost column to 4 in the rightmost column. When you have your total number, divide it by 0.64 to receive your score. (You're dividing by the 16 questions and then multiplying by 25, which adjusts the number to a 100-point scale.) This number is your content hygiene score.

If you fill out the online assessment at https://info.aiim.org/ai-assessment-hygiene-audit we'll do this scoring for you.

# 6. Summary

Within each of the three sections, you received a score. These scores are scaled, and added to an overall number as follows:

| | Personnel & Engagement Score | | Content Access Score | | Content Hygiene Score | | Overall Score |
|---|---|---|---|---|---|---|---|
| ( | | *0.5)+( | | *0.3)+( | | *0.2)= | |

In aggregate, these scores provide a sense for how easy or difficult it will be to implement organizational GenAI projects inside your organization.

**Scores of less than 40 are poor;**

it's likely that implementation of organizational GenAI will be slow, difficult, and prone to failure.

**Scores from 41-70 are normal;**

there is more room for improvement, but the organization has an average likelihood of success with their implementation of GenAI.

**Scores from 71-100 are good,**

indicating better than average chances of a successful implementation.

# 7. About the Contributors

The process of assembling a paper of this magnitude isn't born from one, or two, people. While a few people are acknowledged here, the greater AIIM community has provided a framework for the learning, conversations, and insights that appear here.

## Lead Author:
**Robert Bogue**

**Robert Bogue, CIP,** is the President of Thor Projects LLC and has decades of information management experience, including 21 years as a Microsoft MVP. A bestselling author and international speaker, Robert has enabled thousands to deliver on innovation in corporate, quasi-government, and government organizations. He has authored numerous courses for Microsoft, LinkedIn Learning, Pluralsight, and published Extinguish Burnout: A Practical Guide to Prevention and Recovery, with the Society for Human Resource Management (SHRM). He leverages deep technical expertise, empathetic listening, and business insight to guide customers to success.

Find out more about his consulting work at ThorProjects.com and get tools to be successful in transforming your organization at ConfidentChangeManagement.com.

## Lead Author:
**Tori Miller Liu**

**Tori Miller Liu, MBA, FASAE, CAE, CIP** is the President & CEO of the Association for Intelligent Information Management, the world's leading association dedicated to the information management industry. She is an experienced association executive, technology leader, speaker, and facilitator. Previously, she served as the Chief Information Officer of the American Speech-Language-Hearing Association. A longtime contributor and volunteer in the association community, Tori is a current member of the ASAE Executive Management Advisory Council and a co-founder and past Board Member of the Association Women Technology Champions. She is a Certified Association Executive (CAE), Certified Information Professional (CIP), and holds an MBA from George Washington University. Tori was honored to be named an ASAE Fellow (FASAE) in 2023.

## Contributors

Special thanks and appreciation to Julie Harvey, Amitabh Srivastav, and Jesse Wilkins for their guidance and expertise throughout this project. Without their guidance, this paper wouldn't have been possible. Thank you to Dana Lheureau for her careful editing eye for helping us to translate our thoughts into the prose you see here.

# 8. Resources

The resources that can help you change your scores and to become more prepared for your GenAI success are listed by section below.

## 8.1. Employee Engagement

Enhancing your performance in the ability to execute change in your organization will have the most dramatic influence on your ability to be successful with GenAI.

**A wealth of resources is available at**
https://ConfidentChangeManagement.com/GenerativeAI.

You'll find:

- *A change model library with over 20 different approaches to change*
- *Reviews of over 55 organizational change focused books*
- *Nine resource books consisting of videos and written materials designed to address a specific change management related need*
- *Eight additional assessments of change readiness including psychological safety*
- *Courses on change management*

## 8.2. Content Access and Content Hygiene

Foundational training is available to help information management professionals manage and prepare organizational content for GenAI. Individual courses and comprehensive learning paths are available at aiim.org/training.

You'll find Certificates of Specialization including:

- *Intelligent Information Management Essentials*
- *Modern Records Management*
- *Taxonomy & Metadata*
- *Privacy & Security*
- *and more*

**Additional reading resources are available at**
aiim.org on GenAI.

These include (among others):

*AIIM Guidance on Generative AI for Information Management*

*AIIM Comments on AI Accountability*

*The Critical Role of Content Architecture on Generative AI*

*Data Privacy in the Age of AI*

*Eliminate Bias for Ethical & Responsible Artificial Intelligence*

*and more*


For more information about the CIP, visit
aiim.org/cip

## 8.3. AIIM CIP Certification

AIIM created the Certified Information Professional (CIP) certification to recognize information professionals who have proven they have the readiness skills to protect and manage information and develop organization-wide strategies based on established methodologies and industry best practices. The CIP exam was created by thought leaders in the field of information management based in real-life situations to help CIP credential holders and their organizations stand out in today's information-driven world.

The CIP credential focuses on five critical areas that comprise Intelligent Information Management:

- *Creating, capturing, and sharing information*
- *Extracting intelligence from information*
- *Digitalizing information-intensive processes*
- *Automating governance and compliance*
- *Implementing an information management strategy and solution*

Here is the blueprint containing the full outline of items covered in the current exam.

# 9. Appendix A – Glossary

- ***Application Programming Interface (API)*** *– a software intermediary that allows two applications to talk to each other. APIs are an accessible way to extract and share data within and across organizations.*
- ***Component Object Model (COM)*** *– a legacy Microsoft standard for interprocess communications and API development.*
- ***Content*** *– contains information and addresses a single purpose or goal for the consumer.*
- ***Content Processor*** *– a piece of software designed to extract text out of file formats for consumption by search and AI tools.*
- ***Digital Loss Protection (DLP)*** *– a cybersecurity solution that detects and prevents data breaches.*
- **Encoding** *– the use of code to change original data into a form that can be used by an external process.*
- ***Generative AI*** *– a type of artificial intelligence technology, or algorithms, that can be used to create new high-quality content, including audio, code, images, text, simulations, and videos based on the data they were trained on in response to inputted prompts.*
- ***Markup*** *– a language designed for defining and presenting text (e.g., HTML).*
- ***Metadata*** *– a set of data that describes and gives information about other data.*
- ***Narrow AI*** *– artificial intelligence systems that are designed to perform specific tasks exceptionally well and operate under limited constraints.*
- ***Representational State Transfer (REST)*** *– a software architectural style that defines a common method for defining APIs for web services.*
- ***Simple Objects Access Protocol (SOAP)*** *– an XML-based protocol specification for exchanging structured information to implement web services; enables the distributed elements of an application to communicate.*
- ***Structured Information*** *– information that exists as sets of discrete values within the rows and columns of information tables.*
- ***Taxonomy*** *– a scheme of classification; provides an easy-to-comprehend and navigable visual structure of a knowledge domain by grouping related entities.*
- ***Unstructured Information*** *– information that doesn't fall into rows and columns often documents, presentations, and other non-structured notes.*

# 10. Works Cited

**Association for Intelligent Information Management.**
*(n.d.). 2023 State of the Intelligent Information Management Industry. Retrieved from Association for Intelligent Information Management:*
*https://info.aiim.org/state-of-the-intelligent-information-management-industry-2023*

**Deci, E. L., & Flaste, R. (1995).**
*Why We Do What We Do. New York, NY: Penguin Books.*

**Edmondson, A. (2019).**
*The Fearless Organization: Creating Psychological Safety in the Workplace for Learning, Innovation, and Growth. Hoboken, NJ: John Wiley & Sons.*

**Edmondson, A. (2023).**
*Right Kind of Wrong. New York, NY: Atria Books.*

**Florida, R. (2012).**
*The Rise of the Creative Class, Revisited: Revised and Expanded. New York, NY: Basic Books*.

**Foster, S. (2023, 07 06).**
*Survey: The average American feels the'd need over $200K a year to be financially comfortable. Retrieved 02 29, 2024, from Bankrate:*
*https://www.bankrate.com/personal-finance/financial-freedom-survey/*

**Klein, G. (1998).**
*Sources of Power: How People Make Decisions. Cambridge, MA: The MIT Press.*

**Morgan, G. (2006).**
*Images of Organization. Thousand Oaks, CA: Sage Publications, Inc.*

**Rogers, E. (1962).**
*Diffusion of Innovations. New York, NY: Free Press.*

# About Hyland Software

Hyland

**Hyland Software**
28105 Clemens Road
Westlake, Ohio 44145, USA
+1.440.788.5000
www.hyland.com

Hyland empowers future-focused organizations with AI-powered intelligent content solutions that deliver deep insights so innovators can elevate the experiences of those they serve. Over half of *Fortune* 100 organizations partner with Hyland to unlock the full potential of enterprise content, drive innovation, automate processes, and securely integrate content and data within the business systems they use every day.

Hyland provides the right foundation to help fuel AI success and empower everyday innovators to create better experiences in moments that matter. We bring expertise in turning unstructured data into highly contextualized and AI-ready content to help organizations find information, accelerate decisions and transform the way they work. We also understand the complexity of business processes, and we guide our customers in optimizing the right ones using innovative AI-powered solutions that are tailored for specific industries and business needs, all while being enterprise-class, comprehensive and flexible.

Hyland serves the business needs of enterprise and small-midsize organizations. Our global, 3,000+ employee base serves US/Canada, EMEA, LATAM and APAC across industries and business areas including healthcare, financial services, government, education, insurance, accounts payable, human resources, retail, and manufacturing.

Hyland employees don't do it alone. There are more than 9,000 customers supported through Hyland's global, 4,000 partner community. This group spans across 100 countries. Customers also don't do it alone. Hyland's professional services offer consulting, data conversion, implementation and managed services.

For more information about Hyland, visit
**hyland.com**


aiim
Association for Intelligent Information Management

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.

**Become part of AIIM's vibrant community and join today at**

aiim.org/join

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

aiim
Association for Intelligent Information Management

AVAILTEK