

This report was produced by



Responsibly Sourced Data: AI's Crucial Ingredient

AI winners are focusing on data integrity



Well-managed data gives organizations a strong competitive advantage. Organizations that have superior data collection, storage and analytics capabilities are benefiting from increased revenues and profitability, with use cases ranging from enhanced productivity to new forms of customer experiences and business value propositions.

Now, there is even more potential to get ahead by using artificial intelligence (AI) to drive smarter, faster decision-making. But to avoid problems such as hallucinations, bias and ethical missteps, AI solutions need to be fed with data that is of high quality and responsibly sourced. With the rise of open-source and specialized AI models, data integrity and transparency are more critical than ever.

New research from Iron Mountain, in partnership with FT Longitude, uncovers the extent to which organizations have developed information management systems and datasets that are fit for the AI age. Based on a survey of senior leaders at 500 large organizations worldwide (those with more than 1,000 employees), it identifies the performance gaps in organizations' data ecosystems and practices. And it identifies a group of leaders that are providing a blueprint for building AI-ready information management capabilities.

Defining the leaders

There are 47 organizations in the research - just under 10% of the total respondent base - that demonstrate consistently high performance in converting information management into a competitive advantage.



Key findings

- AI readiness is now seen as critical to organizational ambitions. Respondents in our leader group say that building datasets to power AI is the information management goal that will have the biggest impact on whether they achieve their strategic objectives over the next 12 months
- Large companies have secured a “good data dividend” that is equivalent to average revenue growth of \$1.9bn over the past 12 months
- Most organizations have gaps in their AI readiness. They have managed to achieve occasional benefits from their strategies, but they lack consistency
- The average large organization has lost almost \$390,000 over the past 12 months because of data integrity flaws



How big is the “good data dividend”?

Nine out of 10 organizations in our research have seen both revenue and profitability grow over the past 12 months as a direct result of their information management systems and strategies. Many also report gains in areas including employee productivity, cost management and new business generation.

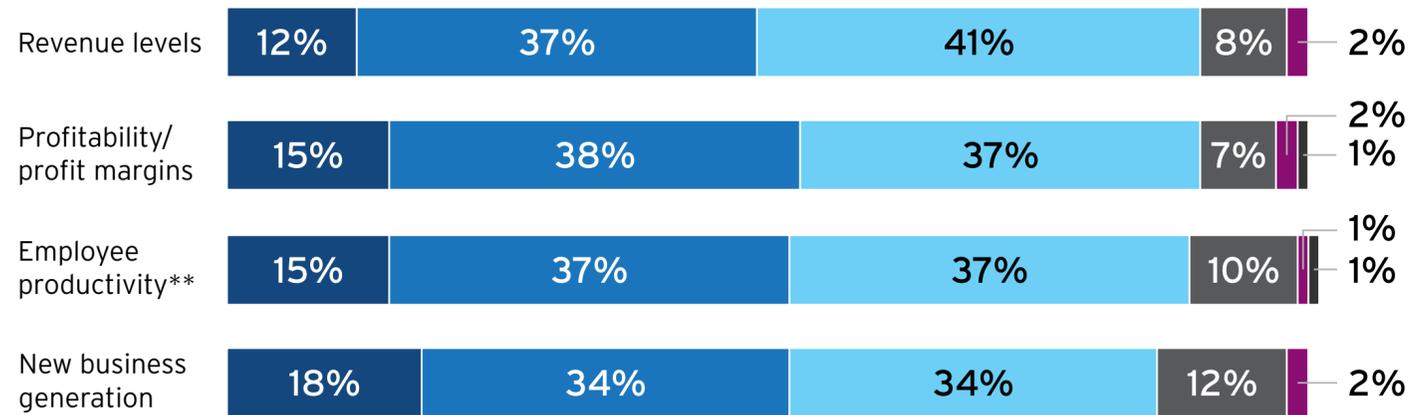
The average organization in our research has secured a “good data dividend” - the increase in revenue due to investment in information management systems and strategies - worth \$1.9bn. This equates to an estimated total global revenue gain of \$72tn.

\$72tn

Large organizations' estimated total global revenue gain from investment in information management systems and strategies

Information management is influencing performance growth

- Increased by more than 20%
- Increased by 10–20%
- Increased by 1–10%
- Stayed the same/no change
- Decreased by 1–10%
- Decreased by 10-20%
- Decreased by more than 20%*



Change in key performance indicators over the past 12 months as a direct result of information management systems and strategies
 *At least 1 respondent provided this answer, but due to rounding its value doesn't reach a percent of the total base



Average revenue gains from information management systems and strategies over the past 12 months

Global total revenue gain:

\$72tn*

Average increase per company in our research:

\$1.9bn

APAC companies:

\$2.9bn

Americas companies:

\$1.9bn

EMEA companies:

\$1.4bn

*full calculation is in the appendix



Organizations are missing out on value

To secure this value, organizations must be ready to make full use of AI. But, most organizations have performance gaps in AI readiness, and the cost of not acting is high, especially when it comes to data integrity, which is a necessity for AI systems.

Organizations say the consequences include reduced agility, poor customer experience and increased operational costs. The average organization in our research estimates it has lost \$389,780 over the past 12 months because of data integrity issues.

About four in 10 organizations say that in areas such as unstructured data, data compliance and data integrity, their activities produce only occasional benefits. The organizations that have consistency flaws in their strategies outnumber the ones whose activities consistently generate value, by 27 percentage points on average.

In the short term, says Narasimha Goli, Chief Technology Officer of Iron Mountain, the way for organizations to increase their success rates is to get started with the information they have rather than seeking perfection by having all their data sources centralized.

“Sometimes, people think they have to centralize all their data before they can start to use it for AI models, but that’s a myth,” says Goli. “Instead, focus on the problem you need to solve and identify the datasets you need to tackle that issue.”

Organizations are under pressure to embrace AI, but most are battling capability gaps. Nearly two-thirds (64%) say their AI readiness activities do not generate consistent value, and 69% say the same about their current levels of data integrity and their ability to source information. These shortfalls could sabotage the reliability of their AI outputs.

Closing these gaps can be challenging: 70% of organizations say they cannot integrate data sources quickly enough for real-time analytics to support their decision-making. And 76% say that technical debt within legacy systems - the cost of reworking systems not fit for current needs - has blocked some of their AI initiatives in the past 12 months. Data quality, cybersecurity and compliance risks, and talent shortages are the biggest hurdles as companies try to make their data infrastructures AI-ready.

Lessons from the leaders

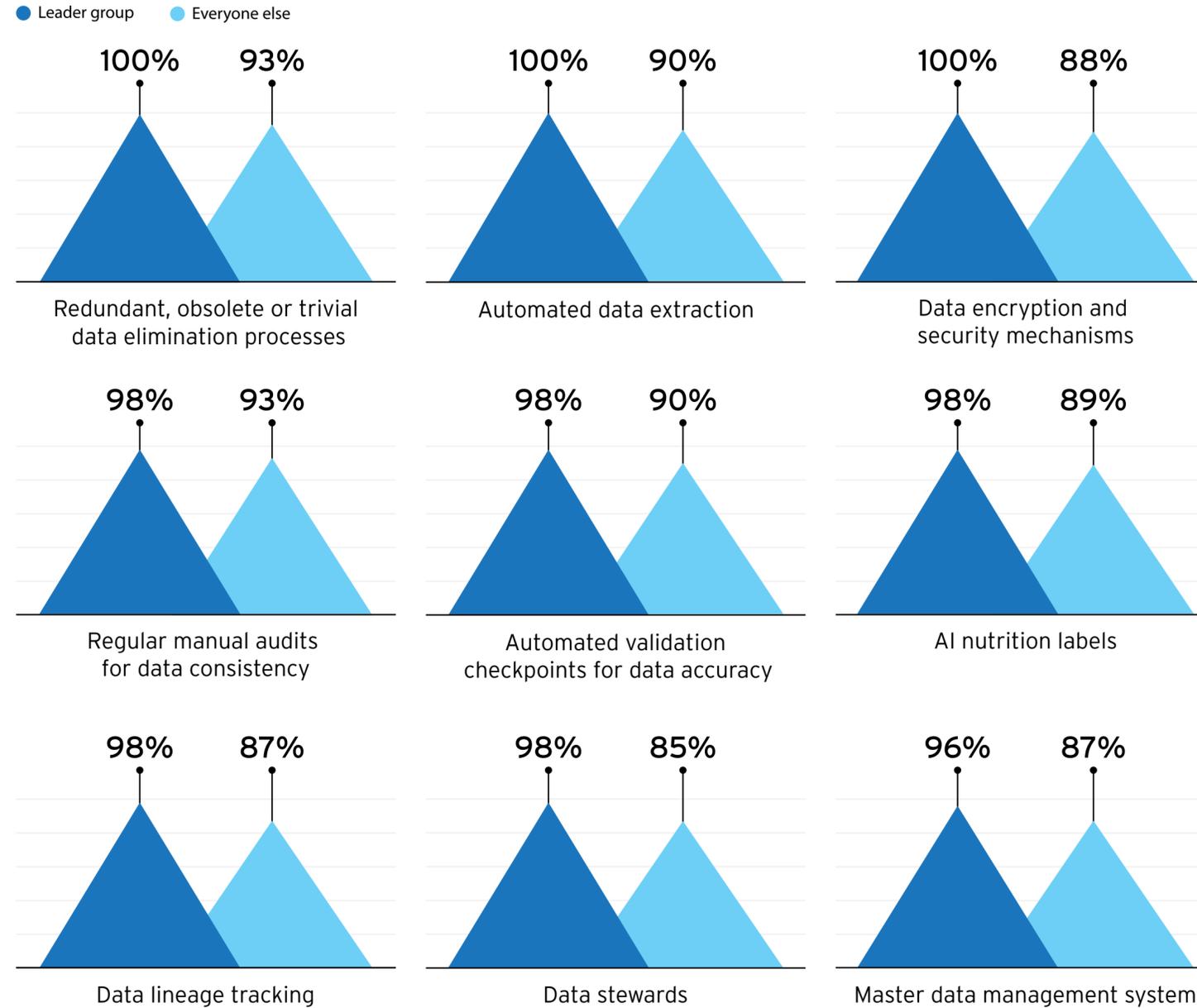
How can these organizations move forward more confidently? Our research identifies a small group of leaders - just under 10% of the organizations in the survey - that are better than others at using information management to create a competitive advantage (greater revenue and profitability increases). These organizations provide vital lessons for the rest.

For example, the leaders have more data integrity and accuracy provisions in their workflows than other organizations. These verification checks will make sure their AI outputs are sourced responsibly.

“AI without good data is like a car without fuel,” says Swami Jayaraman, Senior Vice President and Chief Enterprise Architect at Iron Mountain. “We need to focus on data utility and quality rather than volume and on ensuring that data is gathered with consent.”

Every organization in the leader group has processes for eliminating redundant, obsolete or trivial data and for automating data extraction. And almost all of them carry out regular manual audits for data consistency and have automated validation checkpoints for data accuracy. The leaders are also more likely to prioritize workflow security and embed data encryption alongside other security mechanisms.

The leaders have more data integrity checks in their workflows



“Which of these factors does your organization have in its workflows to manage data integrity and accuracy?”



Find out more about how organizations are paying a price for poor data integrity.



Read more about targeting security and compliance gaps to achieve responsibly sourced AI.

Leaders are more invested in creating a clear line of sight, and AI is helping them

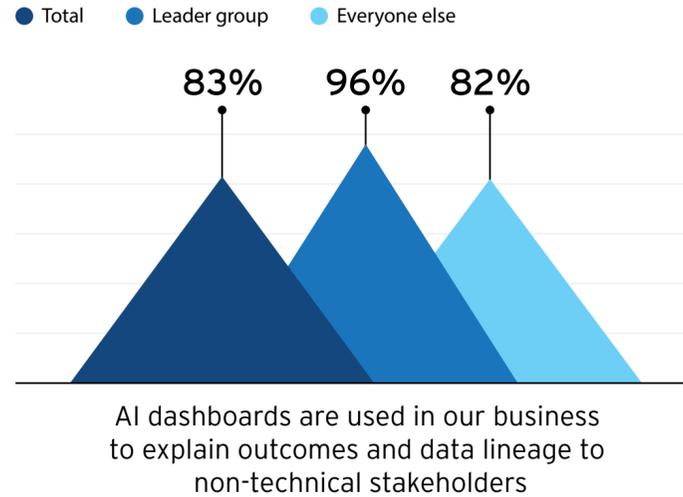
The leaders in the research place more emphasis on data lineage, which is the process of tracking how data is generated, transformed, shared and used across their systems. They are also more likely to use AI dashboards to explain outcomes and data lineage to non-technical stakeholders.

“AI nutrition labels” are another priority for the leaders. Like the nutrition labels on food, these provide important details about the data and processing used by a model. So, they provide transparency about how AI models use data - especially customer data - and build trust in AI.

They can include details on:

- Which AI models are being used
- Whether customer data is used to train the models
- Whether features are optional
- Whether there is “a human in the loop”
- How the data is used, such as for internal projects or to train models for other customers
- Whether personally identifiable information is used

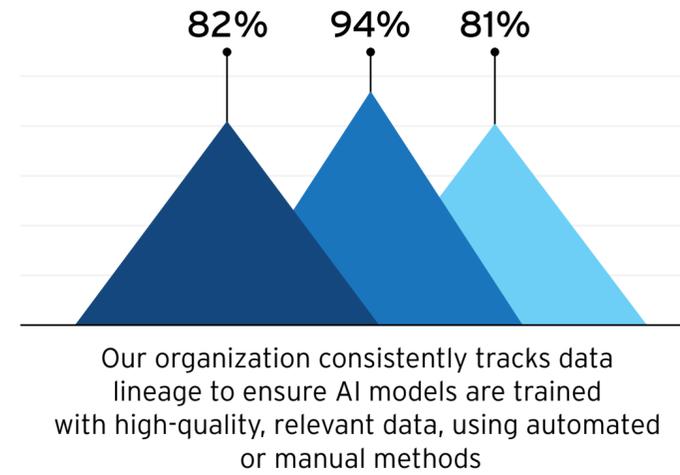
Data lineage for AI is a higher priority for the leaders



The leaders are 16 percentage points more likely than the others to have adopted AI nutrition labels to verify data quality.

So, commitment to responsible data sourcing is a hallmark of the leading organizations. “The way we gain trust is by showing people the data sources we use to create the AI outcomes they see in front of them,” says Jayaraman. “We ground those outcomes in our own datasets, and we provide details of our data lineage to explain where the data comes from.”

Also, the majority of the leaders are using AI technologies to improve their unstructured data sources. Alongside highly skilled personnel and third-party expertise, AI-powered quality control and assurance mechanisms are helping leaders



fill the gaps in their unstructured data sources so they are more AI-ready.

The bottom line, says Goli, is that leaders recognize smart information management is central to capitalizing on the growing AI opportunity. “Before they deploy AI models, or even start to train those models, they focus on the data. They look at where that data is sourced from, how it was received, whether it is tagged and verified, and whether it has holes. With AI fast becoming a necessity for competitive advantage, this data quality-first mindset is now critical for every organization.”



Find out more about how organizations are going full circle with making their unstructured data more AI ready

Appendix: Calculating the total revenue impact of information management

- The following calculations are based on a combination of external data sources, survey findings and applied statistical methods
- The sample focuses only on organizations with more than 1,000 employees
- The majority of companies in the sample have annual revenues between \$1bn and \$99bn

Step 1

For each respondent (organization) in our survey, we identified their approximate annual revenue by using the mid-point of the revenue range they gave us.

Step 2

We identified the average revenue change per organization over the past 12 months due to current information management systems and strategies, also using the mid-point of the range.

Step 3

We multiplied each organization's revenue mid-point by their reported mid-point percentage revenue change. Averaging these values across all respondents in our survey gave us a figure of \$1.913bn per organization.

Step 4

Applying this to an estimated number of organizations worldwide with 1,000+ employees (37,666) gave us the total estimated global revenue change of \$72tn: $37,666 \times \$1,913bn = \sim\$72.07tn$.

What is your organization's total revenue (in US dollars)?

Revenue change	Mid-point
\$10m to \$24.99m	\$18m
\$25m to \$49.99m	\$38m
\$50m to \$99.99m	\$75m
\$100m to \$249.99m	\$175m
\$250m to \$499.99m	\$375m
\$500m to \$999.99m	\$750m
\$1bn to \$9.99bn	\$5.5bn
\$10bn to \$19.99bn	\$15bn
\$20bn to \$49.99bn	\$35bn
\$50bn to \$99.99bn	\$75bn
\$100bn or more	\$200bn

What do you estimate is the change in your organization's revenue levels over the last 12 months as a direct result of your current information management systems and strategies?

Revenue change	Mid-point
Increased by more than 20%	25%
Increased by 10-20%	15%
Increased by 1-10%	5.50%
Stayed the same/no change	0
Decreased by 1-10%	-5.50%
Decreased by 1-10%	-5.50%
Decreased by more than 20%	-25%

This study was based on a survey sample of 500 senior decision-makers. These professionals are at either at C-suite, VP or Director level at companies which have more than 1000 employees. Below are the geographies and sectors that contributed to the survey:

Regions:

Americas: Brazil, Canada and the US.

Europe: Benelux, France, Germany, Italy, Nordics, Spain and the UK.

APAC: Australia, China, Hong Kong, India, Japan, Singapore and South Korea.

Sectors:

Fast moving consumer goods, financial services, healthcare, IT, legal services, life sciences, media and entertainment, oil and gas, pharmaceuticals, public sector, retail, telecoms and utilities.

