Unstructured Data Leads to Increased Risk of Exposure

Adobe's Mroz on Generative AI-Enabled Workflows, New Threats and Layered Defense

ISMG
INFORMATION SECURITY
MEDIA GROUP

Adobe

Geoff Mroz
Director of AI solutions and security

As generative artificial intelligence models continue to accelerate document-driven productivity, it is also expanding the attack surface around unmanaged but sensitive data organizations don't want exposed.

Geoff Mroz, director of AI solutions and security at Adobe, said the list of risks tied to AI-enabled document workflows far outweighs that of traditional file security. Today's defenders are up against sophisticated phishing campaigns, metadata manipulation and inaccurate AI output. As employees adopt unapproved, off-premises tools in the absence of standardized governance rules, the risk of data exposure soars.

"Increased capability also means increased exposure risk … half of organizations out there lack AI governance that's locked into place and their data may be getting exposed in an unmanaged way," Mroz said.

In this video interview with Information Security Media Group, Mroz also discussed:

- How shadow AI is reshaping document security risk;
- Why unstructured data is the most critical AI exposure point;
- Best practices for layered, secure-by-default AI deployments.

As a digital innovator, Mroz influences creative strategy and transformation. With more than 25 years of experience expanding and blurring the lines between art and science, he connect the dots between people, businesses and technology to help companies deliver digital transformation and creative productivity.

"Organizations need layered protection that starts with cloud security, extends to the networks that connect to the cloud and continues to devices such as desktops and servers – applications that run on them and the documents those applications access." – Geof Mroz

DOCUMENT SECURITY RISKS RISE WITH AI

TOM FIELD: Why does document security matter in the AI age and what new vulnerabilities does generative AI introduce into document workflows as organizations adopt it?

GEOFFREY MROZ: Increased capability also means increased exposure risks. Independent research from both 451 Research and IDC confirm that some of the most common risks include things like phishing, metadata manipulation and inaccurate AI output.

SHADOW AI DRIVES UNMANAGED DATA EXPOSURE

FIELD: What are the main factors that lead to these risks?

MROZ: What we're looking at is a variant of shadow IT, which has been around forever. But now, it's more shadow AI, where employees use unapproved tools to boost productivity, probably because most organizations lack AI governance that's locked into place, which increases the risk of unmanaged data exposure.

SECURE BY DEFAULT STRENGTHENS AI WORKFLOWS

FIELD: What best practices support secure and responsible AI and document workflows, and what steps can organizations take to mitigate the risk factors you outlined?

MROZ: Organizations should be looking at secure-by-default vendor solutions as well as things that are not only grounded in responsible AI but have a layered approach to security.

RESPONSIBLE AI ANCHORS SECURE BY DEFAULT

FIELD: Responsible AI and secure by default: What do these mean to you?

MROZ: Responsible AI should be the foundational principle that underpins any AI vendor's commitment to their efforts in developing AI solutions. While ostensibly, this means using AI in responsible ways that accommodate norms and values of our society. What it means is that a vendor needs to ensure that models are trained in a way that is fair, non-discriminating and respects the rights of others, both in terms of bias as well as their intellectual copyrights. Developers should ensure safeguarding user access at each step of the AI generation process, whether it is talking about deny listing certain terms during prompt interpolation or semantic and contextual event filtering and content filtering all the generated results. Developers should ensure that all the solutions are delivered in a way that is not misleading and makes it clear that AI was used in the process. As a good example, at Adobe, we do not train on customer content and we never have. Part of our responsible AI effort is to assure our customers that any data that they give us, we will not be using that data to train our core foundational models.

AI NEEDS DEFENSE ACROSS EVERY LAYER

FIELD: Is success in AI more than just having the right tools?

MROZ: It is about adding a layer of defense and not treating AI as a single feature. Build guardrails and governance into the AI solution itself, from the model to the cloud and the tools and layers that interact with it.

SENSITIVE DATA RESIDES WITHIN UNSTRUCTURED DOCUMENTS

FIELD: What specific layers should organizations consider when building a layered defense?

MROZ: Organizations need layered protection that starts with cloud security, extends to the networks that connect to the cloud and continues to devices such as desktops and servers - applications that run on them and the documents those applications access. Documents in this context refer to unstructured data. Gartner and IDC estimate that 80% to 90% of an organization's data is unstructured. With the rise of large language models and AI, that data presents both an opportunity and a risk. An organization's most sensitive information often resides in unstructured documents and requires protection across every layer to keep data secure.

END-TO-END DOCUMENT SECURITY LAYERS

FIELD: What protections are possible at each of these layers?

MROZ: Let's start with document security. There are things like standard encryption. Documents should be password-protected, either through built-in controls or public key infrastructure-based certificate security. Sensitive information should be redacted, even within internal networks, to reduce exposure if files are shared or leaked. Additional protections include policy enforcement through tools such as Microsoft Purview Information Protection, which helps keep content private. Metadata tagging, including "do not train" labels, can restrict how content is used in LLM retrieval-augmented generation scenarios. These controls help prevent sensitive documents from being ingested or used to train AI systems without authorization.

UNIFIED SECURITY APPROACH LIMITS SHADOW RISKS

FIELD: Is it fair to say this involves significant complexity?

MROZ: Yes, and that's why we recommend that organizations should consolidate the approach and bring everything together. Across these layers of security, the document side is only one part.

There is also application security – how applications are secured and how they handle potentially unsafe information. From a device standpoint, consider licensing, device management and preventing user installs to reduce shadow IT. Address authentication and administrative trust controls across devices organizationwide. From a complexity standpoint, many factors extend to cloud security. What do you enforce at the organizational level to ensure users behave responsibly?

Most users try to do the right thing, but some get lost or go down their own path. How can that be prevented at the organizational level? How is cloud storage encrypted with dedicated keys? How are sharing restrictions enforced so documents cannot be shared outside the organization, even if someone gains access to a file they should not have? Consolidating with a single vendor or fewer vendors and fewer solutions simplifies governance and strengthens the overall security posture.

SECURE BY DEFAULT POWERS ACROBAT PROTECTIONS

FIELD: How is Adobe putting these best practices into action, including secure by default and layered defenses?

MROZ: If you look at Acrobat, which is our core solution for document workflows, and Acrobat AI Assistant, which integrates the power of AI insights and productivity into these flows, these were designed specifically with security guardrails in mind. Most of the layered security are already implemented in Acrobat and the Document Cloud. On the AI side, Acrobat AI Assistant includes document-specific analysis, which include user-defined RAG, short-term memory and 12-hour caching of content. No complete documents are stored in the cloud. The system extracts semantic chunks, stores them temporarily and removes them after 12 hours, which limits exposure and reduces risk. This approach supports insights and automation while minimizing potential data exposure.

PRODUCTIVITY GAINS MUST NOT COMPROMISE CONFIDENTIALITY

FIELD: What should viewers remember most about document protection in the age of AI?

MROZ: Although AI dramatically improves productivity when deployed responsibly, vendor practices should support responsible AI development. Layered security should extend across the organization, starting with documents and continuing to the cloud. This approach enables productivity gains without compromising confidentiality or compliance.

About ISMG

Information Security Media Group (ISMG) is the world's largest media organization devoted solely to information security and risk management. Each of our **38** media properties provides education, research and news that is specifically tailored to key vertical sectors including banking, healthcare and the public sector; geographies from North America to Southeast Asia; and topics such as data breach prevention, cyber risk assessment and fraud. Our annual global summit series connects senior security professionals with industry thought leaders to find actionable solutions for pressing cybersecurity challenges.

(800) 944-0401 • sales@ismg.io

BANK INFO SECURITY® · CU INFO SECURITY® (Just for Credit Unions) · GOV INFO SECURITY® · HEALTHCARE INFO SECURITY®

infoRisk TODAY® · CAREERS INFO SECURITY® · Data Breach TODAY (Prevention. Response. Notification.) · CyberEd.io

CIO.inc · DeviceSecurity.io · PaymentSecurity.io · FraudToday.io

CYBER THEORY · CyberEdBoard · Xtra mile LIFECYCLE MARKETING · GREYHEAD

ISMG
INFORMATION SECURITY
MEDIA GROUP

902 Carnegie Center • Princeton, NJ • 08540 • www.ismg.io